Orthogonal rational functions on the unit circle with prescribed poles not on the unit circle

A. Bultheel, R. Cruz-Barroso, A. Lasarow

Report TW 677, January 2017

KU Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

Orthogonal rational functions on the unit circle with prescribed poles not on the unit circle

A. Bultheel, R. Cruz-Barroso, A. Lasarow

Report TW 677, January 2017

Department of Computer Science, KU Leuven

Abstract

Orthogonal rational functions (ORF) on the unit circle generalize orthogonal polynomials (poles at infinity) and Laurent polynomials (poles at zero and infinity). In this paper we investigate the properties of and the relation between these ORF when the poles are all outside or all inside the unit disk, or when they can be anywhere in the complex plane outside the unit circle.

Keywords : Orthogonal rational functions, recurrence relation, Szegő quadrature, Hessenberg matrix, spectrum
MSC : Primary : 42C10, Secondary : 47A10,65D32

Orthogonal rational functions on the unit circle with prescribed poles not on the unit circle

A. Bultheel [1] R. Cruz-Barroso [2] A. Lasarow [3]

Abstract

Orthogonal rational functions (ORF) on the unit circle generalize orthogonal polynomials (poles at infinity) and Laurent polynomials (poles at zero and infinity). In this paper we investigate the properties of and the relation between these ORF when the poles are all outside or all inside the unit disk, or when they can be anywhere in the complex plane outside the unit circle.

Key words: orthogonal rational functions, Szegő quadrature, spectral method
2000 MSC: 30D15, 30E05, 42C05, 44A60

1 Introduction

Orthogonal rational functions (ORF) on the unit circle are well known as generalizations of orthogonal polynomials on the unit circle (OPUC). The pole at infinity of the polynomials is replaced by poles "in the neighborhood" of infinity, i.e., poles outside the closed unit disk. The recurrence relations for the ORF generalize the Szegő recurrence relations for the polynomials.

If μ is the orthogonality measure supported on the unit circle, and L^2_μ the corresponding Hilbert space, then the shift operator $\mathcal{T}_\mu : L^2_\mu \to L^2_\mu : f(z) \mapsto zf(z)$ restricted to the polynomials has a representation with respect to the orthogonal polynomials that is a Hessenberg matrix. However, if instead of a polynomial basis, one uses a basis or orthogonal Laurent polynomials

[1] Department of Computer Science, KU Leuven, Belgium.
[2] Department of Mathematical Analysis, La Laguna University, Tenerife, Spain.
[3] Fak. Informatik, Mathematik & Naturwissenschaften, HTKW Leipzig, Germany

by alternating between poles at infinity and poles at the origin, a full unitary representation of $\mathcal{T}_\mu$ with respect to this basis is a five-diagonal CMV matrix [6].

The previous ideas have been generalized to the rational case by Velázquez in [28]. He showed that the representation of the shift operator with respect to the classical ORF is not a Hessenberg matrix but a matrix Möbius transform of a Hessenberg matrix. However, a full unitary representation can be obtained if the shift is represented with respect to a rational analog of the Laurent polynomials by alternating between a pole inside and a pole outside the unit disk. The resulting matrix is a matrix Möbius transform of a five-diagonal matrix.

Orthogonal Laurent polynomials on the real line, a half-line, or an interval were introduced by Jones et al. [18,19] in the context of moment problems, Padé approximation and quadrature and this was elaborated by many authors. González-Vera and his coworkers were in particular involved in extending the theory where the poles zero and infinity alternate (the so-called balanced situation) to a more general case where in each step either infinity or zero can be chosen as a pole in any arbitrary order [11,4]. They also identify the resulting orthogonal Laurent polynomials as shifted versions of the orthogonal polynomials. Hence the orthogonal Laurent polynomials satisfy the same recurrence as the classical orthogonal polynomials after an appropriate shifting and normalization is embedded.

The corresponding case of orthogonal Laurent polynomials on the unit circle were introduced by Thron in [25] and have been studied more recently in for example [7,9]. Papers traditionally deal with the balanced situation like in [9] but in [7] also an arbitrary ordering was considered. Only in [8] Cruz-Barroso and Delvaux investigated the structure of the matrix representation with respect to the basis of the resulting orthogonal Laurent polynomials on the circle. They called it a "snake-shaped" matrix which generalizes the five diagonal matrix.

The purpose of this paper is to generalize these ideas valid for Laurent polynomials on the circle to the rational case. That is to choose the poles of the ORF in an arbitrary order either inside or outside the unit disk. We relate the resulting ORF with the ORF having all their poles outside or all their poles inside the disk, and study the corresponding recurrence relations. With respect to this new orthogonal rational basis, the shift operator will be represented by a matrix Möbius transformation of a snake-shaped matrix.

In the papers by Lasarow and coworkers (e.g. [20,14,16,15]) matrix versions of the ORF are considered. In these papers also an arbitrary choice of the poles is allowed but only with the restrictive condition that if α is used as a pole, then $1/\overline{\alpha}$ cannot be used anymore. This means that for example the "balanced situation" is excluded. One of the goals of this paper is to remove this restriction on the poles.

In the context of quadrature formulas, an arbitrary sequence of poles not on the unit circle was also briefly discussed in [10]. The sequence of poles considered there need not be Newtonian, i.e., the poles for the ORF of degree n may depend on n. Since our approach will emphasize the role of the recurrence relation for the ORF, we do need a Newtonian sequence, although some of the results may be generalizable to the situation of a non-Newtonian sequence of poles.

One of the applications of the theory of ORF is the construction of quadrature formulas on the unit circle that form rational generalizations of the Szegő quadrature. They are exact in spaces of rational functions having poles inside and outside the unit disk. The nodes of the quadrature formula are zeros of para-orthogonal rational functions (PORF) and the weights are all positive numbers. These nodes and weights can (like in Gaussian quadrature) be derived from the eigenvalue decomposition of a unitary truncation of the shift operator to a finite dimensional subspace.

The outline of the paper is as follows. In Section 2 we introduce the main notations used in this paper. The linear spaces and the ORF bases are given in Section 3. Section 4 brings the Christoffel-Darboux relations and the reproducing kernels which form an essential element to obtain the recurrence relation given in Section 5 but also for the PORF in Section 6 to be used for quadrature formulas in Section 7. The alternative representation of the shift operator is given in Section 8 and its factorization in elementary 2×2 blocks in the subsequent Section 9. We end by drawing some conclusions about the spectrum of the shift operator and about the computation of rational Szegő quadrature formulas in sections 10 and 11. The ideas that we have presented in this paper, especially the factorization of unitary Hessenberg matrices in elementary unitary factors is also used in the linear algebra literature in the finite dimensional situation. The algorithms are quite similar but not the same as we explain briefly in Section 12. For the details one will have to look up the references.

2 Basic definitions and notation

We use the following notation. $\mathbb{C}$ denotes the complex plane, $\hat{\mathbb{C}}$ the extended complex plane (one point compactification), $\mathbb{R}$ the real line, $\hat{\mathbb{R}}$ the closure of $\mathbb{R}$ in $\hat{\mathbb{C}}$, $\mathbb{T}$ the unit circle, $\mathbb{D}$ the open unit disk, $\hat{\mathbb{D}} = \mathbb{D} \cup \mathbb{T}$, and $\mathbb{E} = \hat{\mathbb{C}} \setminus \hat{\mathbb{D}}$. For any number $z \in \hat{\mathbb{C}}$ we define $z_* = 1/\overline{z}$ (and set $1/0 = \infty$, $1/\infty = 0$) and for any complex function f, we define $f_*(z) = \overline{f(z_*)}$.

To approximate an integral

$$I_\mu(f) = \int_{\mathbb{T}} f(z)\,\mathrm{d}\mu(z)$$

where μ is a probability measure on $\mathbb{T}$ one may use Szegő quadrature formulas. The nodes of this quadrature can be computed by using the Szegő polynomials. Orthogonality in this paper will always be with respect to the inner product

$$\langle f, g\rangle = \int_{\mathbb{T}} \overline{f(z)} g(z) d\mu(z).$$

The weights of the n-point quadrature are all positive and the formula is exact for all Laurent polynomials $f \in \operatorname{span}\{z^k : |k| \leq n-1\}$.

This has been generalized to rational functions with a set of predefined poles. The corresponding quadrature formulas are then rational Szegő quadratures. See for example [5]. The idea is the following. Fix a sequence $\underline{\boldsymbol{\alpha}} = \{\alpha_1, \alpha_2, \ldots\} \subset \mathbb{D}$, and consider the subspaces of rational functions defined by

$$\mathcal{L}_0 = \mathbb{C}, \quad \mathcal{L}_n = \left\{ \frac{p_n(z)}{\pi_n(z)} : p_n \in \mathcal{P}_n, \quad \pi_n(z) = \prod_{k=1}^{n} (1 - \overline{\alpha}_k z) \right\}, \quad n \geq 1$$

where $\mathcal{P}_n$ is the set of polynomials of degree at most n. These rational functions have their poles among the points in $\underline{\boldsymbol{\alpha}}_* = \{\alpha_{j*} = 1/\overline{\alpha}_j : \alpha_j \in \underline{\boldsymbol{\alpha}}\}$. Let $\phi_n \in \mathcal{L}_n \setminus \mathcal{L}_{n-1}$, and $\phi_n \perp \mathcal{L}_{n-1}$ be the nth orthogonal rational basis function (ORF) in a nested sequence. These functions have all their zeros in $\mathbb{D}$, while the quadrature formulas one has in mind should have their nodes on the circle $\mathbb{T}$. Therefore, para-orthogonal rational functions (PORF) are introduced defined by

$$Q_n(z, \tau) = \phi_n(z) + \tau \phi_n^*(z), \quad \tau \in \mathbb{T}$$

where besides the ORF $\phi_n(z) = \frac{p_n(z)}{\pi_n(z)}$ also the "reciprocal" function $\phi_n^*(z) = \frac{p_n^*(z)}{\pi_n(z)} = \frac{z^n p_{n*}(z)}{\pi_n(z)}$, is introduced. These PORF have n simple zeros $\{\xi_{nk}\}_{k=1}^n \subset \mathbb{T}$ so that they can be used as nodes for the quadrature formulas

$$I_n(f) = \sum_{k=1}^{n} \lambda_{nk} f(\xi_{nk})$$

and the weights are all positive, given by $\lambda_{nk} = 1/\sum_{k=0}^{n-1} |\phi_k(\xi_{nk})|^2$. These quadrature formulas are exact for all functions of the form $\{f = g_* h : g, h \in \mathcal{L}_{n-1}\} = \mathcal{L}_{n-1}\mathcal{L}_{(n-1)*}$.

The purpose of this paper is to generalize the situation where the α_j are all in $\mathbb{D}$ to the situation where they are anywhere in the extended complex plane outside $\mathbb{T}$. This will require the introduction of some new notation.

So consider a sequence $\underline{\boldsymbol{\alpha}} \subset \mathbb{D}$ and its reflection in the circle $\underline{\boldsymbol{\beta}} \subset \mathbb{E}$ where $\beta_j = 1/\overline{\alpha}_j = \alpha_{j*}$ for all $j = 1, 2, \ldots$. We now construct a new sequence $\underline{\boldsymbol{\gamma}} = \{\gamma_1, \gamma_2, \ldots\}$ where each γ_j is either equal to α_j or β_j.

Partition $\{1,2,\ldots,n\}$ ($n \in \hat{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$) into two disjoint index sets: the ones where $\gamma_j = \alpha_j$ and the indices where $\gamma_j = \beta_j$:

$$\mathbb{a}_n = \{j : \gamma_j = \alpha_j \in \mathbb{D},\ 1 \le j \le n\} \quad \text{and} \quad \mathbb{b}_n = \{j : \gamma_j = \beta_j \in \mathbb{E},\ 1 \le j \le n\}.$$

and define

$$\underline{\boldsymbol{\alpha}}_n = \{\alpha_j : j \in \mathbb{a}_n\} \quad \text{and} \quad \underline{\boldsymbol{\beta}}_n = \{\beta_j : j \in \mathbb{b}_n\}.$$

Occasionally it will be useful to prepend the sequence $\underline{\boldsymbol{\alpha}}$ with an extra point $\alpha_0 = 0$. That means that $\underline{\boldsymbol{\beta}}$ is preceded by $\beta_0 = 1/\overline{\alpha}_0 = \infty$. For $\underline{\boldsymbol{\gamma}}$, the initial point can be $\gamma_0 = \alpha_0 = 0$ or $\gamma_0 = \beta_0 = \infty$.

With each of the series $\underline{\boldsymbol{\alpha}}$, $\underline{\boldsymbol{\beta}}$, and $\underline{\boldsymbol{\gamma}}$ we can associate orthogonal rational functions. They will be closely related as we shall show. They will differ only by a Blaschke product factor just like the orthogonal Laurent polynomials are essentially shifted versions of the orthogonal polynomials [7].

To define the denominators of our rational functions, we introduce the following elementary factors

$$\varpi_j^\alpha(z) = 1 - \overline{\alpha}_j z, \quad \varpi_j^\beta(z) = \begin{cases} 1 - \overline{\beta}_j z, \text{ if } \beta_j \neq \infty, \\ -z, \qquad\quad \text{if } \beta_j = \infty. \end{cases} \quad \varpi_j^\gamma(z) = \begin{cases} \varpi_j^\alpha(z), \text{ if } \gamma_j = \alpha_n, \\ \varpi_j^\beta(z), \text{ if } \gamma_j = \beta_n. \end{cases}$$

Note that if $\alpha_j = 0$ then $\varpi_j^\alpha(z) = 1$ and hence $\beta_j = \infty$ but $\varpi_j^\beta(z) = -z$.

To separate the α and the β-factors in a product, we also define

$$\dot{\varpi}_j^\alpha(z) = \begin{cases} \varpi_j^\alpha, \text{ if } \gamma_j = \alpha_j \\ 1, \quad \text{if } \gamma_j = \beta_j \end{cases}, \quad \text{and} \quad \dot{\varpi}_j^\beta(z) = \begin{cases} \varpi_j^\beta, \text{ if } \gamma_j = \beta_j \\ 1, \quad \text{if } \gamma_j = \alpha_j. \end{cases}$$

Because the sequence $\underline{\boldsymbol{\gamma}}$ is our main focus, we simplify the notation by removing the superscript γ when not needed. E.g. $\varpi_j(z) = \varpi_j^\gamma(z) = \dot{\varpi}_j^\alpha \dot{\varpi}_j^\beta$ etc.

We can now define for $\nu \in \{\alpha, \beta, \gamma\}$

$$\pi_n^\nu(z) = \prod_{j=1}^n \varpi_j^\nu(z)$$

and the reduced products separating the α and the β-factors

$$\dot{\pi}_n^\alpha(z) = \prod_{j=1}^n \dot{\varpi}_j^\alpha(z) = \prod_{j\in \mathbb{a}_n} \varpi_j(z), \quad \dot{\pi}_n^\beta(z) = \prod_{j=1}^n \dot{\varpi}_j^\beta(z) = \prod_{j\in \mathbb{b}_n} \varpi_j(z)$$

so that

$$\pi_n(z) = \prod_{j=1}^n \varpi_j(z) = \dot{\pi}_n^\alpha(z)\dot{\pi}_n^\beta(z).$$

We assume here and in the rest of the paper that products over $j \in \varnothing$ equal 1.

The Blaschke factors are defined for $\nu \in \{\alpha, \beta, \gamma\}$ as

$$\begin{aligned}
\zeta_j^\nu(z) &= \sigma_j^\nu \frac{z-\nu_j}{1-\overline{\nu}_j z}, \quad \sigma_j^\nu = \tfrac{\overline{\nu}_j}{|\nu_j|}, \text{ if } \nu_j \notin \{0,\infty\},\\
\zeta_j^\nu(z) &= \sigma_j^\nu z = z, \qquad \sigma_j^\nu = 1, \quad \text{if } \nu_j = 0,\\
\zeta_j^\nu(z) &= \sigma_j^\nu/z = 1/z, \ \sigma_j^\nu = 1, \quad \text{if } \nu_j = \infty.
\end{aligned}$$

Thus

$$\sigma_j^\nu = \begin{cases} \frac{\overline{\nu}_j}{|\nu_j|} \text{ for } & \nu_j \notin \{0,\infty\}, \\ 1, \text{ for } & \nu_j \in \{0,\infty\}. \end{cases}$$

Because $\sigma_n^\alpha = \sigma_n^\beta$, we can remove the superscript and just write σ_n.

We shall also use the following notation which maps complex numbers onto $\mathbb{T}$

$$\mathfrak{u}(z) = \begin{cases} \frac{\overline{z}}{|z|} \in \mathbb{T}, \forall z \in \mathbb{C}\setminus\{0\} \\ 1, \qquad\quad z \in \{0,\infty\} \end{cases}$$

then $\sigma_j = \mathfrak{u}(\alpha_j) = \mathfrak{u}(\beta_j) = \mathfrak{u}(\nu_j)$.

Set $(\varpi_j^\nu)^*(z) = \varpi_j^{\nu*}(z) = z\varpi_{j*}^\nu(z)$ (i.e. $(1-\overline{\alpha}_j z)^* = z - \alpha_j$ if e.g. $\nu = \alpha$), then $\zeta_j^\nu = \sigma_j \frac{\varpi_j^{\nu*}}{\varpi_j^\nu}$. Later we shall also use $\pi_n^{\nu*}$ to mean $\prod_{j=1}^n \varpi_j^{\nu*}$. Note that $\zeta_j^\alpha = \zeta_{j*}^\beta = 1/\zeta_j^\beta$. Moreover if $\alpha_j = 0$ and hence $\beta_j = \infty$, then $\varpi_j^{\alpha*}(z) = z$ and $\varpi_j^{\beta*}(z) = -1$.

Next define the finite Blaschke products for $\nu \in \{\alpha, \beta\}$

$$B_0^\nu = 1, \quad \text{and} \quad B_n^\nu(z) = \prod_{j=1}^n \zeta_j^\nu(z), \quad n = 1, 2, \ldots.$$

It is important to note that here $\nu \neq \gamma$. For the definition of $B_n^\gamma = B_n$ see below.

Like we have split up the denominators $\pi_n = \dot{\pi}_n^\alpha \dot{\pi}_n^\beta$ in the α-factors and the β-factors, we define for $n \geq 1$

$$\dot{\zeta}_j^\alpha = \begin{cases} \zeta_j^\alpha, \text{ if } & \gamma_j = \alpha_j \\ 1, \quad \text{if } & \gamma_j = \beta_j \end{cases}, \quad \dot{\zeta}_j^\beta = \begin{cases} \zeta_j^\beta, \text{ if } & \gamma_j = \beta_j \\ 1, \quad \text{if } & \gamma_j = \alpha_j \end{cases}$$

and

$$\dot{B}_n^\alpha(z) = \prod_{j=1}^n \dot{\zeta}_j^\alpha(z) = \prod_{j \in \mathbb{a}_n} \zeta_j(z), \quad \text{and} \quad \dot{B}_n^\beta(z) = \prod_{j=1}^n \dot{\zeta}_j^\beta(z) = \prod_{j \in \mathbb{b}_n} \zeta_j(z)$$

so that we can define the finite Blaschke products for the $\underline{\gamma}$

$$B_n(z) = \begin{cases} \dot{B}_n^\alpha(z), \text{ if } & \gamma_n = \alpha_n \\ \dot{B}_n^\beta(z), \text{ if } & \gamma_n = \beta_n. \end{cases}$$

Note that the reflection property of the factors also holds for the products: $B_n^\alpha = (B_n^\beta)_* = 1/B_n^\beta$. However $\dot{B}_n^\alpha = \prod_{j \in \mathbb{a}_n} \zeta_j^\alpha = \prod_{j \in \mathbb{a}_n} \zeta_{j*}^\beta = \prod_{j \in \mathbb{a}_n} 1/\zeta_j^\beta \neq \prod_{j \in \mathbb{b}_n} 1/\zeta_j^\beta = 1/\dot{B}_n^\beta$ and thus also $B_{n*} \neq 1/B_n$. However $(\dot{B}_n^\alpha \dot{B}_n^\beta)_* = 1/(\dot{B}_n^\beta \dot{B}_n^\alpha)$.

3 Linear spaces and ORF bases

We can now introduce our spaces of rational functions for $n \geq 0$:

$$\mathcal{L}_n^\nu = \operatorname{span}\{B_0^\nu, B_1^\nu, \ldots, B_n^\nu\}, \quad \nu \in \{\alpha, \beta, \gamma\}, \quad \text{and} \quad \dot{\mathcal{L}}_n^\nu = \operatorname{span}\{\dot{B}_0^\nu, \dot{B}_1^\nu, \ldots, \dot{B}_n^\nu\}, \quad \nu \in \{\alpha, \beta\}.$$

The dimension of $\mathcal{L}_n^\nu$ is $n+1$ for $\nu \in \{\alpha, \beta, \gamma\}$, but note that the dimension of $\dot{\mathcal{L}}_n^\nu$ for $\nu \in \{\alpha, \beta\}$ can be less than $n+1$. Indeed some of the $\dot{B}_j^\nu$ may be repeated so that for example the dimension of $\dot{\mathcal{L}}_n^\alpha$ is only $|\mathbb{a}_n| + 1$ with $|\mathbb{a}_n|$ the cardinality of $\mathbb{a}_n$ and similarly for $\nu = \beta$. Hence for $\nu = \gamma$:

$$\mathcal{L}_n = \operatorname{span}\{B_0, \ldots, B_n\} = \operatorname{span}\{\dot{B}_0, \dot{B}_1^\alpha, \ldots, \dot{B}_n^\alpha, \dot{B}_1^\beta, \ldots, \dot{B}_n^\beta\} = \dot{\mathcal{L}}_n^\alpha \cup \dot{\mathcal{L}}_n^\beta.$$

Because for $n \geq 1$

$$\dot{B}_n^\alpha = \prod_{j \in \mathbb{a}_n} \zeta_j^\alpha = \prod_{j \in \mathbb{a}_n} \frac{1}{\zeta_j^\beta} \quad \text{and} \quad \dot{B}_n^\beta = \prod_{j \in \mathbb{b}_n} \zeta_j^\beta = \prod_{j \in \mathbb{b}_n} \frac{1}{\zeta_j^\alpha},$$

it should be clear that $B_k^\alpha = \dot{B}_k^\alpha/\dot{B}_k^\beta$ and $B_k^\beta = \dot{B}_k^\beta/\dot{B}_k^\alpha$, hence that

$$\mathcal{L}_n^\alpha = \operatorname{span}\left\{\dot{B}_0, \frac{\dot{B}_1^\alpha}{\dot{B}_1^\beta}, \ldots, \frac{\dot{B}_n^\alpha}{\dot{B}_n^\beta}\right\} \quad \text{and} \quad \mathcal{L}_n^\beta = \operatorname{span}\left\{\dot{B}_0, \frac{\dot{B}_1^\beta}{\dot{B}_1^\alpha}, \ldots, \frac{\dot{B}_n^\beta}{\dot{B}_n^\alpha}\right\}.$$

Occasionally we shall also need the notation

$$\dot{\varsigma}_n^\alpha = \prod_{j\in \mathbb{a}_n} \sigma_j \in \mathbb{T}, \quad \dot{\varsigma}_n^\beta = \prod_{j\in \mathbb{b}_n} \sigma_j \in \mathbb{T}, \quad \text{and} \quad \varsigma_n = \prod_{j=1}^n \sigma_j \in \mathbb{T}.$$

Lemma 3.1 *If $f \in \mathcal{L}_n$ then $f/\dot{B}_n^\beta \in \mathcal{L}_n^\alpha$ and $f/\dot{B}_n^\alpha \in \mathcal{L}_n^\beta$. In other words $\mathcal{L}_n = \dot{B}_n^\beta\mathcal{L}_n^\alpha = \dot{B}_n^\alpha\mathcal{L}_n^\beta$. This is true for all $n \geq 0$ is we set $\dot{B}_0^\alpha = \dot{B}_0^\beta = 1$.*

PROOF. This is trivial for $n = 0$ since then $\mathcal{L}_n = \mathbb{C}$.
If $f \in \mathcal{L}_n$, and $n \geq 1$ then it is of the form

$$f(z) = \frac{p_n(z)}{\pi_n(z)} = \frac{p_n(z)}{\dot{\pi}_n^\alpha(z)\dot{\pi}_n^\beta(z)}, \quad p_n \in \mathcal{P}_n.$$

Therefore

$$\frac{f(z)}{\dot{B}_n^\beta(z)} = \overline{\dot{\varsigma}}_n^\beta \frac{p_n(z)\dot{\pi}_n^\beta(z)}{\dot{\pi}_n^\alpha(z)\dot{\pi}_n^\beta(z)\dot{\pi}_n^{\beta*}(z)} = \overline{\dot{\varsigma}}_n^\beta \frac{p_n(z)}{\dot{\pi}_n^\alpha(z)\dot{\pi}_n^{\beta*}(z)}.$$

Recall that $\varpi_j^{\beta*} = -1$ and $\sigma_j = 1$ if $\beta_j = \infty$ (and hence $\alpha_j = 0$), we can leave these factors out and we shall write $\dot{\prod}$ for the product instead of $\prod$, the dot meaning that we leave out all the factors for which $\alpha_j = 1/\overline{\beta}_j = 0$.

$$\frac{\overline{\dot{\varsigma}}_n^\beta}{\dot{\pi}_n^{\beta*}(z)} = \dot{\prod_{j\in\mathbb{b}_n}} \frac{\beta_j}{|\beta_j|(z-\beta_j)} = \dot{\prod_{j\in\mathbb{b}_n}} \frac{|\alpha_j|}{\overline{\alpha}_j(z-1/\overline{\alpha}_j)} = \dot{\prod_{j\in\mathbb{b}_n}} \frac{-|\alpha_j|}{1-\overline{\alpha}_j z},$$

and thus

$$\frac{f(z)}{\dot{B}_n^\beta(z)} = c_n \frac{p_n(z)}{\prod_{j=1}^n (1-\overline{\alpha}_j z)} \in \mathcal{L}_n^\alpha, \quad c_n = \dot{\prod_{j\in\mathbb{b}_n}} (-|\alpha_j|) \neq 0.$$

The second part is similar. $\square$

Lemma 3.2 *With our previous definitions we have for $n \geq 1$*

$$\dot{B}_n^\beta\mathcal{L}_{n-1}^\alpha = \operatorname{span}\left\{B_k^\alpha\dot{B}_n^\beta = \frac{\dot{B}_k^\alpha}{\dot{B}_k^\beta}\dot{B}_n^\beta : k = 0, \ldots, n-1\right\} = \dot{\zeta}_n^\beta \operatorname{span}\{B_0, B_1, \ldots, B_{n-1}\} = \dot{\zeta}_n^\beta\mathcal{L}_{n-1},$$

and similarly

$$\dot{B}_n^\alpha \mathcal{L}_{n-1}^\beta = \operatorname{span}\left\{ B_k^\beta \dot{B}_n^\alpha = \frac{\dot{B}_k^\beta}{\dot{B}_k^\alpha} \dot{B}_n^\alpha : k = 0, \dots, n-1 \right\} = \dot{\zeta}_n^\alpha \operatorname{span}\{B_0, B_1, \dots, B_{n-1}\} = \dot{\zeta}_n^\alpha \mathcal{L}_{n-1}.$$

PROOF. By our previous lemma $\dot{B}_n^\beta \mathcal{L}_{n-1}^\alpha = \dot{\zeta}_n^\beta \dot{B}_{n-1}^\beta \mathcal{L}_{n-1}^\alpha = \dot{\zeta}_n^\beta \mathcal{L}_{n-1}$. The second relation is proved in a similar way. $\square$

To introduce the sequences of orthogonal rational functions (ORF) for the different sequences $\underline{\boldsymbol{\nu}}$, $\nu \in \{\alpha, \beta, \gamma\}$ recall the inner product that we can write with our $_*$-notation as $\langle f, g \rangle = \int_{\mathbb{T}} f_*(z) g(z)\, \mathrm{d}\mu(z)$ where μ is assumed to be a probability measure positive a.e. on $\mathbb{T}$.

Then the orthogonal rational functions (ORF) with respect to the sequence $\underline{\boldsymbol{\nu}}$ with $\nu \in \{\alpha, \beta, \gamma\}$ are defined by $\phi_n^\nu \in \mathcal{L}_n^\nu$ with $\phi_n^\nu \perp \mathcal{L}_{n-1}^\nu$ for $n \geq 1$ and we choose $\phi_0^\nu = 1$.

Lemma 3.3 *The function $\phi_n^\alpha \dot{B}_n^\beta$ belongs to $\mathcal{L}_n$ and it is orthogonal to the $(n-1)$-dimensional subspace $\dot{\zeta}_n^\beta \mathcal{L}_{n-1}$ for all $n \geq 1$.*

Similarly, the function $\phi_n^\beta \dot{B}_n^\alpha$ belongs to $\mathcal{L}_n$ and it is orthogonal to the $(n-1)$-dimensional subspace $\dot{\zeta}_n^\alpha \mathcal{L}_{n-1}$, $n \geq 1$.

PROOF. First note that $\phi_n^\alpha \dot{B}_n^\beta \in \mathcal{L}_n$ by Lemma 3.1.

By definition $\phi_n^\alpha \perp \mathcal{L}_{n-1}^\alpha$. Thus by Lemma 3.2 and because $\langle f, g \rangle = \left\langle \dot{B}_n^\nu f, \dot{B}_n^\nu g \right\rangle$,

$$\dot{B}_n^\beta \phi_n^\alpha \perp \dot{B}_n^\beta \mathcal{L}_{n-1}^\alpha = \dot{\zeta}_n^\beta \mathcal{L}_{n-1}.$$

The second claim follows by symmetry. $\square$

Note that $\dot{\zeta}_n^\beta \mathcal{L}_{n-1} = \mathcal{L}_{n-1}$ if $\gamma_n = \alpha_n$. Thus, up to normalization, $\phi_n^\alpha \dot{B}_n^\beta$ is the same as ϕ_n and similarly, if $\gamma_n = \beta_n$ then ϕ_n and $\phi_n^\beta \dot{B}_n^\alpha$ are the same up to normalization.

Lemma 3.4 *For $n \geq 1$ the function $\dot{B}_n^\alpha (\phi_n^\alpha)_*$ belongs to $\mathcal{L}_n$ and it is orthogonal to $\dot{\zeta}_n^\alpha \mathcal{L}_{n-1}$.*

Similarly, for $n \geq 1$ the function $\dot{B}_n^\beta (\phi_n^\beta)_*$ belongs to $\mathcal{L}_n$ and it is orthogonal to $\dot{\zeta}_n^\beta \mathcal{L}_{n-1}$.

PROOF. Since $\phi_n^\alpha \dot{B}_n^\beta \perp \dot{\zeta}_n^\beta \mathcal{L}_{n-1}$,

$$(\phi_n^\alpha \dot{B}_n^\beta)_* \perp \dot{\zeta}_{n*}^\beta \mathcal{L}_{(n-1)*},$$

and thus by Lemma 3.2 and because

$$\dot{B}_{n-1}^\alpha \dot{B}_{n-1}^\beta \mathcal{L}_{(n-1)*} = \dot{B}_{n-1}^\alpha \dot{B}_{n-1}^\beta \frac{\mathcal{P}_{(n-1)*}}{\dot{\pi}_{(n-1)*}^\alpha \dot{\pi}_{(n-1)*}^\beta} = \frac{\mathcal{P}_{n-1}}{\dot{\pi}_{n-1}^\alpha \dot{\pi}_{n-1}^\beta} = \mathcal{L}_{n-1}$$

it follows that

$$\dot{B}_n^\alpha \phi_{n*}^\alpha = \dot{B}_n^\alpha \dot{B}_n^\beta (\phi_n^\alpha \dot{B}_n^\beta)_* \perp \dot{\zeta}_n^\alpha \dot{B}_{n-1}^\alpha \dot{B}_{n-1}^\beta \mathcal{L}_{(n-1)*} = \dot{\zeta}_n^\alpha \mathcal{L}_{n-1}.$$

The other claim follows by symmetry. $\square$

We now define the reciprocal ORFs by (recall $f_*(z) = \overline{f(1/\overline{z})}$)

$$(\phi_n^\nu)^* = B_n^\nu (\phi_n^\nu)_* \quad \nu \in \{\alpha, \beta\}.$$

For the ORF in $\mathcal{L}_n$ however we set

$$\phi_n^* = \dot{B}_n^\alpha \dot{B}_n^\beta (\phi_n)_*.$$

Note that by definition B_n is either $\dot{B}_n^\alpha$ or $\dot{B}_n^\beta$ depending on γ_n being α_n or β_n, while in the previous definition we do not multiply with B_n but with the product $\dot{B}_n^\alpha \dot{B}_n^\beta$. The reason is that we want the operation $(\ )^*$ to be a map from $\mathcal{L}_n^\nu$ to $\mathcal{L}_n^\nu$.

Remark 3.5 As the operation $(\ )^*$ is a map from $\mathcal{L}_n^\nu$ to $\mathcal{L}_n^\nu$, it depends on n and on ν. So to make the notation unambiguous we should in fact use something like $f^{[\nu,n]}$ if $f \in \mathcal{L}_n^\nu$. However, in order not to overload our notation, we shall stick to the notation f^* since it should always be clear from the context what the space is to which f will belong. Note that we also used the same notation to transform polynomials. That is conform our agreement because a polynomial of degree n belongs to $\mathcal{L}_n^\alpha$ for a sequence $\underline{\boldsymbol{\alpha}}$ where all $\alpha_j = 0$, $j = 1, 2, \ldots$ and for this sequence $B_n^\alpha(z) = z^n$.

The orthogonality conditions define ϕ_n and ϕ_n^* uniquely up to normalization.

Suppose $\gamma_n = \alpha_n$, then ϕ_n and $\phi_n^\alpha \dot{B}_n^\beta$ are both in $\mathcal{L}_n$ and orthogonal to $\mathcal{L}_{n-1}$ (Lemma 3.3). If we normalize ϕ_n^ν then from $\|\phi_n\| = 1$ and $\|\phi_n^\alpha \dot{B}_n^\beta\| = \|\phi_n^\alpha\| = 1$, it follows that there must be some unimodular constant $s_n^\alpha \in \mathbb{T}$ such that $\phi_n = s_n^\alpha \phi_n^\alpha \dot{B}_n^\beta$. Of course, we have by symmetry that for $\gamma_n = \beta_n$, there is some $s_n^\beta \in \mathbb{T}$ such that $\phi_n = s_n^\beta \phi_n^\beta \dot{B}_n^\alpha$.

To fix the unimodular factors s_n^α and s_n^β, we first normalize ϕ_n^α and ϕ_n^β as follows.

For ϕ_n^α we know that $\phi_n^{\alpha*}(\alpha_n) \neq 0$ because all its zeros are in $\mathbb{E}$. So that as a normalization we can take $\phi_n^{\alpha*}(\alpha_n) > 0$. Similarly for ϕ_n^β we can normalize by $\phi_n^{\beta*}(\beta_n) > 0$. In both cases, we have made the leading coefficient with respect to the basis $\{B_j^\nu\}_{j=0}^n$ positive since $\phi_n^\alpha(z) = \overline{\phi_n^{\alpha*}(\alpha_n)}B_n^\alpha(z) + \psi_{n-1}^\alpha(z)$ with $\psi_{n-1}^\alpha \in \mathcal{L}_{n-1}^\alpha$ and $\phi_n^\beta(z) = \overline{\phi_n^{\beta*}(\beta_n)}B_n^\beta(z) + \psi_{n-1}^\beta(z)$ with $\psi_{n-1}^\beta \in \mathcal{L}_{n-1}^\beta$

Before we define the normalization for the $\boldsymbol{\gamma}$-sequence, we prove the following Lemma which is a consequence of the normalization of the ϕ_n^α and the ϕ_n^β.

Lemma 3.6 *For the orthonormal ORFs, it holds that $\phi_n^\alpha = \phi_{n*}^\beta$ and $(\phi_n^\alpha)^*\dot{B}_n^\beta = \phi_n^\beta \dot{B}_n^\alpha$ and hence also $(\phi_n^\beta)^*\dot{B}_n^\alpha = \phi_n^\alpha \dot{B}_n^\beta$ for all $n \geq 0$.*

PROOF. For $n = 0$, this is trivial since $\phi_0, \phi_0^\alpha, \phi_0^\beta, \dot{B}_0^\alpha$ and $\dot{B}_0^\beta$ are all equal to 1.
We give the proof for $n \geq 1$ and $\gamma_n = \alpha_n$ (for $\gamma_n = \beta_n$, the proof is similar). Since by previous lemmas $\dot{B}_n^\beta(\phi_n^\beta)_*$ and $\phi_n^\alpha \dot{B}_n^\beta$ are both in $\mathcal{L}_n$ and orthogonal to $\mathcal{L}_{n-1}$, and since $\|\dot{B}_n^\beta(\phi_n^\beta)_*\| = \|\phi_n^\beta\| = 1$ and $\|\phi_n^\alpha \dot{B}_n^\beta\| = \|\phi_n^\alpha\| = 1$, there must be some $s_n \in \mathbb{T}$ such that

$$s_n\phi_n^\alpha \dot{B}_n^\beta = \phi_{n*}^\beta \dot{B}_n^\beta \quad \text{or} \quad s_n\phi_n^\alpha = \phi_{n*}^\beta.$$

Multiply with $B_n^\beta = B_{n*}^\alpha$ and evaluate at β_n to get $s_n\phi_n^\alpha(\beta_n)B_{n*}^\alpha(\beta_n) = \phi_n^{\beta*}(\beta_n) > 0$. Thus s_n should arrange for

$$0 < s_n\phi_n^\alpha(1/\overline{\alpha}_n)B_{n*}^\alpha(1/\overline{\alpha}_n) = s_n\overline{\phi_{n*}^\alpha(\alpha_n)B_n^\alpha(\alpha_n)} = s_n\overline{\phi_n^{\alpha*}(\alpha_n)},$$

and since $\phi_n^{\alpha*}(\alpha_n) > 0$, it follows that $s_n = 1$.

Because $(\phi_n^\alpha)^* = B_n^\alpha\phi_{n*}^\alpha = B_n^\alpha\phi_n^\beta$ and $B_n^\alpha = \dot{B}_n^\alpha/\dot{B}_n^\beta$, also the other claims follow. $\square$

For the normalization of the ϕ_n, we can do two things: either we make the normalization of ϕ_n simple and choose for example $\phi_n^*(\gamma_n) > 0$, similar to what we did for ϕ_n^α and ϕ_n^β (but this is somewhat problematic as we shall see below), or we can insist on keeping the relation with ϕ_n^α and ϕ_n^β simple as in the previous lemma and arrange that $s_n^\alpha = s_n^\beta = 1$. We choose for the second option.

Let us assume that $\gamma_n = \alpha_n$. Denote

$$\phi_n(z) = \frac{p_n(z)}{\dot{\pi}_n^\alpha(z)\dot{\pi}_n^\beta(z)} \quad \text{and} \quad \phi_n^\alpha(z) = \frac{p_n^\alpha(z)}{\pi_n^\alpha(z)},$$

with p_n and p_n^α both polynomials in $\mathcal{P}_n$. Then

$$\phi_n^*(z) = \frac{\varsigma_n\, p_n^*(z)}{\dot\pi_n^\alpha(z)\dot\pi_n^\beta(z)} \quad \text{and} \quad \phi_n^{\alpha*}(z) = \frac{\varsigma_n\, p_n^{\alpha*}(z)}{\pi_n^\alpha(z)}, \quad \varsigma_n = \prod_{j=1}^n \sigma_j.$$

We already know that there is some $s_n^\alpha \in \mathbb{T}$ such that $\phi_n = s_n^\alpha \dot B_n^\beta \phi_n^\alpha$. Take the $(\;)_*$ conjugate and multiply with $\dot B_n^\alpha \dot B_n^\beta$ to get $\phi_n^* = \overline{s}_n^\alpha \dot B_n^\beta \phi_n^{\alpha*}$.

It now takes some simple algebra to reformulate $\phi_n^* = \overline{s}_n^\alpha \dot B_n^\beta \phi_n^{\alpha*}$ as

$$\phi_n^*(z) = \frac{\varsigma_n\, p_n^*(z)}{\dot\pi_n^\alpha(z)\dot\pi_n^\beta(z)} = \overline{s_n^\alpha} \frac{\varsigma_n\, p_n^{\alpha*}(z)}{\dot\pi_n^\alpha(z)\dot\pi_n^\beta(z)} \dot\prod_{j\in\mathbb{b}_n} (-|\beta_j|).$$

This implies that $p_n^*(z) = \overline{s_n^\alpha} p_n^{\alpha*}(z) \dot\prod_{j\in\mathbb{b}_n}(-|\beta_j|)$ and thus that $p_n^*(z)$ has the same zeros as $p_n^{\alpha*}(z)$, none of which is in $\mathbb{D}$. Thus the numerator of ϕ_n^* will not vanish at $\alpha_n \in \mathbb{D}$ but one of the factors $(1-\overline{\beta}_j\alpha_n)$ from $\dot\pi_n^\beta(\alpha_n)$ could be zero. Thus a normalization $\phi_n^*(\alpha_n) > 0$ is not an option in general. We could however make $s_n^\alpha = 1$ when we choose $p_n^*(\alpha_n)/p_n^{\alpha*}(\alpha_n) > 0$ or, since $\phi_n^{\alpha*}(\alpha_n) > 0$, this is equivalent with $\varsigma_n\, p_n^*(\alpha_n)/\pi_n^\alpha(\alpha_n) > 0$. Yet another way to put this is requiring that $\phi_n^*(z)/\dot B_n^\beta(z)$ is positive at $z = \alpha_n$. This does not give a problem with 0 or ∞ since

$$\dot B_n^\alpha(z)\phi_{n*}(z) = \frac{\phi_n^*(z)}{\dot B_n^\beta(z)} = \frac{\dot\varsigma_n^\alpha\, p_n^*(z)}{\dot\pi_n^\alpha(z)\dot\prod_{j\in\mathbb{b}_n}(z-\beta_j)}, \quad \dot\varsigma_n^\alpha = \prod_{j\in\mathbb{a}_n} \sigma_j. \tag{3.1}$$

It is clear that neither the numerator nor the denominator will vanish for $z = \alpha_n$.

Of course a similar argument can be given if $\gamma_n = \beta_n$. Then we choose $\phi_n^*(z)/\dot B_n^\alpha(z)$ to be positive at $z = \beta_n$ or equivalently $\varsigma_n\, p_n^*(\beta_n)/\pi_n^\beta(\beta_n)\dot\prod_{j\in\mathbb{a}_n}(-|\alpha_j|) > 0$.

Let us formulate the result about the numerators as a lemma for further reference.

Lemma 3.7 *With the normalization that we just imposed the numerators p_n^ν of $\phi_n^\nu = p_n^\nu/\pi_n^\nu$, $\nu \in \{\alpha, \beta, \gamma\}$ and $n \geq 1$ are related by*

$$p_n(z) = p_n^\alpha(z) \dot\prod_{j\in\mathbb{b}_n} (-|\beta_j|) = p_n^{\beta*}(z)\varsigma_n \dot\prod_{j\in\mathbb{a}_n} (-|\alpha_j|), \quad \textit{if} \;\; \gamma_n = \alpha_n$$

and

$$p_n(z) = p_n^\beta(z) \dot\prod_{j\in\mathbb{a}_n} (-|\alpha_j|) = p_n^{\alpha*}(z)\varsigma_n \dot\prod_{j\in\mathbb{b}_n} (-|\beta_j|), \quad \textit{if} \;\; \gamma_n = \beta_n$$

where as before $\varsigma_n = \prod_{j=1}^n \sigma_j$.

PROOF. The first expression for $\gamma_n = \alpha_n$ has been proved above. The second one follows in a similar way form the relation $\phi_n(z) = \phi_n^{\beta*}(z)\dot{B}_n^{\alpha}(z)$. Indeed

$$\begin{aligned}\frac{p_n(z)}{\pi_n(z)} &= \frac{\varsigma_n p_n^{\beta*}(z)}{\prod_{j\in\mathbb{a}_n}\varpi_j^{\beta}(z)\prod_{j\in\mathbb{b}_n}\varpi_j^{\beta}(z)}\prod_{j\in\mathbb{a}_n}\sigma_j\frac{z-\alpha_j}{1-\overline{\alpha}_j z}\\ &= \frac{\varsigma_n p_n^{\beta*}(z)}{\prod_{j\in\mathbb{a}_n}\varpi_j^{\alpha}(z)\prod_{j\in\mathbb{b}_n}\varpi_j^{\beta}(z)}\prod_{j\in\mathbb{a}_n}\sigma_j\frac{z-\alpha_j}{1-\overline{\beta}_j z} = \frac{\varsigma_n p_n^{\beta*}(z)}{\pi_n(z)}\prod_{j\in\mathbb{a}_n}\sigma_j(-\alpha_j)\frac{z-\alpha_j}{z-\alpha_j}.\end{aligned}$$

With $-\sigma_j\alpha_j = -|\alpha_j|$ the result follows.

The case $\gamma_n = \beta_n$ is similar. $\square$

Note that this normalization does again mean that we take the leading coefficient of ϕ_n to be positive in the following sense. If $\gamma_n = \alpha_n$ then $\phi_n(z) = \overline{(\dot{B}_n^{\alpha}\phi_{n*})(\alpha_n)}\dot{B}_n^{\alpha}(z)+\psi_{n-1}(z)$ with $\psi_{n-1}\in\mathcal{L}_{n-1}$. Since $\dot{B}_n^{\alpha}\phi_{n*} = \phi_n^{\alpha*}$ and $\phi_n^{\alpha*}(\alpha_n) > 0$. If $\gamma_n = \beta_n$ then $\phi_n(z) = \overline{(\dot{B}_n^{\beta}\phi_{n*})(\beta_n)}\dot{B}_n^{\beta}(z)+\psi_{n-1}(z)$ with $\psi_{n-1}\in\mathcal{L}_{n-1}$ and the conclusion follows similarly.

Whenever we use the term orthonormal, we assume this normalization.

Thus we have proved the following Theorem. It says that if $\gamma_n = \alpha_n$, then ϕ_n is a 'shifted' version of ϕ_n^{α} where 'shifted' means multiplied by $\dot{B}_n^{\beta}$:

$$\dot{B}_n^{\beta}(z)\phi_n(z) = \dot{B}_n^{\beta}(z)[a_0B_0^{\alpha}+\cdots+a_nB_n^{\alpha}(z)] = a_0\dot{B}_n^{\beta}(z)+\cdots+a_n\dot{B}_n^{\alpha}(z),$$

and a similar interpretation if $\gamma_n = \beta_n$. We summarize this in the following theorem.

Theorem 3.8 *If all ORFs ϕ_n^{ν}, $\nu\in\{\alpha,\beta,\gamma\}$ are orthonormal with positive leading coefficient, i.e.,*

$$\phi_n^{\alpha*}(\alpha_n) > 0 \quad \textit{and} \quad \phi_n^{\beta*}(\beta_n) > 0 \quad \textit{and} \quad \begin{cases}(\phi_n^*/\dot{B}_n^{\beta})(\alpha_n) > 0 & \textit{if} \ \ \gamma_n = \alpha_n\\ (\phi_n^*/\dot{B}_n^{\alpha})(\beta_n) > 0 & \textit{if} \ \ \gamma_n = \beta_n.\end{cases}$$

Then for all $n \geq 0$

$$\phi_n = (\phi_n^{\alpha})\dot{B}_n^{\beta} = (\phi_n^{\beta})^*\dot{B}_n^{\alpha} \quad \textit{and} \quad \phi_n^* = (\phi_n^{\alpha})^*\dot{B}_n^{\beta} = (\phi_n^{\beta})\dot{B}_n^{\alpha} \quad \textit{if}\ \gamma_n = \alpha_n$$

while

$$\phi_n = (\phi_n^{\beta})\dot{B}_n^{\alpha} = (\phi_n^{\alpha})^*\dot{B}_n^{\beta} \quad \textit{and} \quad \phi_n^* = (\phi_n^{\beta})^*\dot{B}_n^{\alpha} = (\phi_n^{\alpha})\dot{B}_n^{\beta} \quad \textit{if}\ \gamma_n = \beta_n.$$

Corollary 3.9 *We have for all $n \geq 1$ that $(\phi_n^\nu)^* \perp \zeta_n^\nu \mathcal{L}_{n-1}^\nu$, $\nu \in \{\alpha, \beta, \gamma\}$.*

Corollary 3.10 *The rational functions ϕ_n^α and $\phi_n^{\alpha *}$ are in $\mathcal{L}_n^\alpha$ and hence have all their poles in $\{\beta_j : j = 1, \ldots, n\} \subset \mathbb{E}$ while the zeros of ϕ_n^α are all in $\mathbb{D}$ and the zeros of $\phi_n^{\alpha *}$ are all in $\mathbb{E}$.*
The rational functions ϕ_n^β and $\phi_n^{\beta *}$ are in $\mathcal{L}_n^\beta$ and hence have all their poles in $\{\alpha_j : j = 1, \ldots, n\} \subset \mathbb{D}$ while the zeros of ϕ_n^β are all in $\mathbb{E}$ and the zeros of $\phi_n^{\beta *}$ are all in $\mathbb{D}$.
The rational functions ϕ_n and ϕ_n^* are in $\mathcal{L}_n$ and hence have all their poles in $\{\beta_j : j \in \mathbb{a}_n\} \cup \{\alpha_j : j \in \mathbb{b}_n\}$.
The zeros of ϕ_n are the same as the zeros of ϕ_n^α and thus are all in $\mathbb{D}$ if $\gamma_n = \alpha_n$ and they are the same as the zeros of ϕ_n^β and thus they are all in $\mathbb{E}$ if $\gamma_n = \beta_n$.

PROOF. It is well known that the zeros of ϕ_n^α are all in $\mathbb{D}$, and because $\phi_n^\beta = \phi_{n*}^\alpha$, this means that the zeros of ϕ_n^β are all in $\mathbb{E}$.

Because $\phi_n = (\phi_n^\alpha)\dot{B}_n^\beta = (\phi_n^\alpha)/\prod_{j \in \mathbb{b}_n} \zeta_j^\alpha$ if $\gamma_n = \alpha_n$, i.e., $n \in \mathbb{a}_n$, and the product with $\dot{B}_n^\beta$ will only exchange the poles $1/\overline{\alpha}_j = \beta_j$, $j \in \mathbb{b}_n$ in ϕ_n for poles $\alpha_j = 1/\overline{\beta}_j$, the zeros of ϕ_n^α are left unaltered.

The proof for $n \in \mathbb{b}_n$ is similar. $\square$

One may summarize that for $f \in \mathcal{L}_n^\nu$ the f_* transform reflects both zeros and poles in $\mathbb{T}$ since $z \mapsto z_* = 1/\overline{z}$, while the transform $f \to f^*$ as it is defined in the spaces $\mathcal{L}_n^\nu$, $\nu \in \{\alpha, \beta, \gamma\}$, keeps the poles but reflects the zeros since the multiplication with the respective factors B_n^α, B_n^β and $\dot{B}_n^\alpha \dot{B}_n^\beta$ will only undo the reflection of the poles that resulted from the f_* operation.

4 Christoffel-Darboux relations and reproducing kernels

For $\nu \in \{\alpha, \beta, \gamma\}$, one may define the reproducing kernels for the space $\mathcal{L}_n^\nu$. Given the ORF ϕ_k^ν, the kernels are defined by

$$k_n^\nu(z, w) = \sum_{k=0}^{n} \phi_n^\nu(z)\overline{\phi_n^\nu(w)}.$$

They reproduce $f \in \mathcal{L}_n^\nu$ by $\langle k_n^\nu(\cdot, z), f \rangle = f(z)$.

The proof of the Christoffel-Darboux relations goes exactly like in the classical case and we shall not repeat it here. See e.g. [5, Theorem 3.1.3].

Theorem 4.1 *The Christoffel-Darboux relations*

$$k_n^\nu(z,w) = \frac{\phi_n^{\nu*}(z)\overline{\phi_n^{\nu*}(w)} - \zeta_n^\nu(z)\overline{\zeta_n^\nu(w)}\phi_n^\nu(z)\overline{\phi_n^\nu(w)}}{1-\zeta_n^\nu(z)\overline{\zeta_n^\nu(w)}} = \frac{\phi_{n+1}^{\nu*}(z)\overline{\phi_{n+1}^{\nu*}(w)} - \phi_{n+1}^\nu(z)\overline{\phi_{n+1}^\nu(w)}}{1-\zeta_{n+1}^\nu(z)\overline{\zeta_{n+1}^\nu(w)}}$$

hold for $n \geq 0$, $\nu \in \{\alpha, \beta, \gamma\}$ *and* z, w *not among the poles of* ϕ_n^ν *and not on* $\mathbb{T}$.

As an immediate consequence we have:

Theorem 4.2 *The following relations hold true:*

$$k_n^\alpha(z,w)\dot{B}_n^\beta(z)\overline{\dot{B}_n^\beta(w)} = k_n(z,w) = k_n^\beta(z,w)\dot{B}_n^\alpha(z)\overline{\dot{B}_n^\alpha(w)}$$

for $n \geq 0$ *and* $z, w \notin (\mathbb{T} \cup \{\beta_j : j \in \mathbb{a}_n\} \cup \{\alpha_j : j \in \mathbb{b}_n\})$.

PROOF. The first relation was directly shown above for the case $\gamma_n = \alpha_n$. It also follows in the case $\gamma_{n+1} = \alpha_{n+1}$ and using in the second CD-relation the first expressions from Theorem 3.8 for ϕ_{n+1} and ϕ_{n+1}^*. The relation is thus valid independent of $\gamma_n = \alpha_n$ or $\gamma_n = \beta_n$.

Similarly the second expression was derived before in the case $\gamma_n = \beta_n$, but again, it also follows from the second CD-relation and the first expressions from Theorem 3.8 for ϕ_{n+1} that ϕ_{n+1}^* in the case $\gamma_{n+1} = \beta_{n+1}$. Again the relation holds independently of $\gamma_n = \alpha_n$ or $\gamma_n = \beta_n$.

Alternatively, the second relation can also be derived from the second CD-relation in the case $\gamma_{n+1} = \alpha_{n+1}$ but using the second expressions from Theorem 3.8 for ϕ_{n+1} and ϕ_{n+1}^*. $\square$

Evaluation of the CD-relation in ν_n for $\nu \in \{\alpha, \beta\}$ results in another useful corollary.

Corollary 4.3 *For* $\nu \in \{\alpha, \beta\}$ *we have for* $n \geq 0$

$$k_n^\nu(z,\nu_n) = \phi_n^{\nu*}(z)\overline{\phi_n^{\nu*}(\nu_n)} \quad \textit{and} \quad k_n^\nu(\nu_n,\nu_n) = |\phi_n^{\nu*}(\nu_n)|^2.$$

The latter corollary cannot immediately be used when $\nu = \gamma$ because γ_n could be equal to some pole of ϕ_n if it equals some $1/\overline{\gamma}_j$ for $j < n$. In that case we can remove the denominators in the CD relation and only keep the numerators. Hence setting

$$k_n(z,w) = \frac{P_n(z,w)}{\pi_n(z)\overline{\pi_n(w)}}, \quad \phi_n(z) = \frac{p_n(z)}{\pi_n(z)}, \quad \phi_n^*(z) = \frac{\varsigma_n p_n^*(z)}{\pi_n(z)}, \quad \varsigma_n \in \mathbb{T},$$

the CD relation becomes

$$P_n(z,w) = \frac{p_n^*(z)\overline{p_n^*(w)} - \zeta_n(z)\overline{\zeta_n(w)}p_n(z)\overline{p_n(w)}}{1-\zeta_n(z)\overline{\zeta_n(w)}} = \frac{p_{n+1}^*(z)\overline{p_{n+1}^*(w)} - p_{n+1}(z)\overline{p_{n+1}(w)}}{(1-\zeta_{n+1}(z)\overline{\zeta_{n+1}(w)})\varpi_{n+1}(z)\overline{\varpi_{n+1}(w)}}. \tag{4.1}$$

Thus, the first form gives

$$P_n(z,\gamma_n) = p_n^*(z)\overline{p_n^*(\gamma_n)} \quad \text{and} \quad P_n(\gamma_n,\gamma_n) = |p_n^*(\gamma_n)|^2.$$

Evaluating a polynomial at infinity means taking its highest degree coefficient, i.e., if $q_n(z)$ is supposed to be a polynomial of degree n, then $q_n(\infty)$ stands for its coefficient of z^n.

The second form of (4.1) gives for $\gamma_{n+1} \neq \infty$ and $\gamma_n \neq \infty$

$$P_n(z,\gamma_n) = \frac{p_{n+1}^*(z)\overline{p_{n+1}^*(\gamma_n)} - p_{n+1}(z)\overline{p_{n+1}(\gamma_n)}}{(1-\overline{\gamma}_n z)(1-|\gamma_{n+1}|^2)} \quad \text{and} \quad P_n(\gamma_n,\gamma_n) = \frac{|p_{n+1}^*(\gamma_n)|^2 - |p_{n+1}(\gamma_n)|^2}{(1-|\gamma_n|^2)(1-|\gamma_{n+1}|^2)}.$$

Coupling the first and the second form in (4.1) gives

$$\frac{|p_{n+1}^*(\gamma_n)|^2 - |p_{n+1}(\gamma_n)|^2}{(1-|\gamma_n|^2)(1-|\gamma_{n+1}|^2)} = |p_n^*(\gamma_n)|^2.$$

For $\gamma_{n+1} = \infty$ and $\gamma_n \neq \infty$ we get

$$P_n(z,\gamma_n) = \frac{p_{n+1}^*(z)\overline{p_{n+1}^*(\gamma_n)} - p_{n+1}(z)\overline{p_{n+1}(\gamma_n)}}{-(1-\overline{\gamma}_n z)} = p_n^*(z)\overline{p_n^*(\gamma_n)}$$

and

$$P_n(\gamma_n,\gamma_n) = \frac{|p_{n+1}^*(\gamma_n)|^2 - |p_{n+1}(\gamma_n)|^2}{-(1-|\gamma_n|^2)} = |p_n^*(\gamma_n)|^2.$$

If $\gamma_{n+1} = \infty$ and $\gamma_n = \infty$, the denominators in (4.1) have to be replaced by 1, which gives

$$P_n(z,\gamma_n) = p_{n+1}^*(z)\overline{p_{n+1}^*(\gamma_n)} - p_{n+1}(z)\overline{p_{n+1}(\gamma_n)} = p_n^*(z)\overline{p_n^*(\gamma_n)}$$

and

$$P_n(\gamma_n,\gamma_n) = |p_{n+1}^*(\gamma_n)|^2 - |p_{n+1}(\gamma_n)|^2 = |p_n^*(\gamma_n)|^2.$$

For $\gamma_n = \infty$ and $\gamma_{n+1} \neq \infty$ we obtain in a similar way

$$P_n(z,\gamma_n) = \frac{p_{n+1}^*(z)\overline{p_{n+1}^*(\gamma_n)} - p_{n+1}(z)\overline{p_{n+1}(\gamma_n)}}{z(1-|\gamma_{n+1}|^2)} = p_n^*(z)\overline{p_n^*(\gamma_n)}$$

and

$$P_n(\gamma_n,\gamma_n) = \frac{|p^*_{n+1}(\gamma_n)|^2 - |p_{n+1}(\gamma_n)|^2}{(1-|\gamma_{n+1}|^2)} = |p^*_n(\gamma_n)|^2.$$

To summarize, the relations of Corollary 4.3 may not hold for the ORF if $\nu=\gamma$, but similar relations do hold for the numerators as stated in the next corollary.

Corollary 4.4 *If $P_n(z,w)$ is the numerator in the CD relation and $p_n(z)$ is the numerator of the ORF ϕ_n for the sequence $\underline{\boldsymbol{\gamma}}$ then we have for $n \geq 0$*

$$P_n(z,\gamma_n) = p^*_n(z)\overline{p^*_n(\gamma_n)} \quad \textit{and} \quad P_n(\gamma_n,\gamma_n) = |p^*_n(\gamma_n)|^2.$$

5 Recurrence relation

The recurrence for the ϕ^α_n is well known. For a proof see e.g., [5]. For ϕ^β_n the proof can be copied by symmetry. However, also for $\nu=\gamma$ the same recurrence and its proof can be copied, with this exception that the derivation fails when $p^*_n(\gamma_{n-1})=0$ where $p_n=\phi_n\pi_n$. This can (only) happen if $(1-|\gamma_n|)(1-|\gamma_{n-1}|)<0$ (i.e., one of these γ's is in $\mathbb{D}$ and the other is in $\mathbb{E}$). We shall say that ϕ_n is *regular* if $p^*_n(\gamma_{n-1})\neq 0$. If $\nu=\alpha$ or $\nu=\beta$ then the whole sequence $(\phi^\nu_n)_{n\geq 0}$ will be automatically regular. Thus we have the following theorem:

Theorem 5.1 *Let $\nu\in\{\alpha,\beta,\gamma\}$ and if $\nu=\gamma$ assume moreover that ϕ^ν_n is regular, then the following recursion holds with initial condition $\phi^\nu_0=\phi^{\nu*}_0=1$*

$$\begin{bmatrix}\phi^\nu_n(z)\\ \phi^{\nu*}_n(z)\end{bmatrix} = N^\nu_n \frac{\varpi^\nu_{n-1}(z)}{\varpi^\nu_n(z)} \begin{bmatrix}1 & \lambda^\nu_n\\ \overline{\lambda}^\nu_n & 1\end{bmatrix} \begin{bmatrix}\zeta^\nu_{n-1}(z) & 0\\ 0 & 1\end{bmatrix} \begin{bmatrix}\phi^\nu_{n-1}(z)\\ \phi^{\nu*}_{n-1}(z)\end{bmatrix},$$

where N^ν_n is a nonzero constant times a unitary matrix:

$$N^\nu_n = e^\nu_n \begin{bmatrix}\eta^\nu_{n1} & 0\\ 0 & \eta^\nu_{n2}\end{bmatrix}, \quad e^\nu_n\in\mathbb{C}\setminus\{0\}, \quad \eta^\nu_{n1},\eta^\nu_{n2}\in\mathbb{T}.$$

The constant η^ν_{n1} is chosen such that the normalization condition for the ORFs is maintained. The other constant η^ν_{n2} is then automatically related to η^ν_{n1} by $\eta^\nu_{n2}=\overline{\eta}^\nu_{n1}\overline{\sigma}_{n-1}\sigma_n$. The Szegő parameter λ^ν_n is given by

$$\lambda^\nu_n = \eta^\nu_n \frac{p^\nu_n(\nu_{n-1})}{p^{\nu*}_n(\nu_{n-1})} \quad \textit{with} \quad \eta^\nu_n = \overline{\varsigma}_{n-2}\frac{\overline{p^{\nu*}_{n-1}(\nu_{n-1})}}{p^{\nu*}_{n-1}(\nu_{n-1})} \in\mathbb{T}$$

where $\phi_n^\nu(z) = p_n^\nu(z)/\pi_n^\nu(z)$.

PROOF. The proof for $\nu = \alpha$ is well known. See e.g., [5, Thm. 4.1.1]. For $\nu = \beta$, the same proof holds by symmetry. For $\nu = \gamma$ we repeat the proof to see in detail what happens. For simplicity assume that γ_n and γ_{n-1} are not 0 and not ∞. The technicalities when this is not true are left as an exercise. It is easy because formally it follows the steps of the proof below but one has to replacing a linear factor with infinity by the coefficient of infinity (like $1-\infty z = -z$ and $z - \infty = -1$) and evaluating a polynomial at ∞ means taking its leading coefficient.

First we show that there are some numbers c_n and d_n such that

$$\phi(z) = \frac{1-\overline{\gamma}_n z}{z-\gamma_{n-1}}\phi_n(z) - d_n\phi_{n-1}(z) - c_n\frac{1-\overline{\gamma}_{n-1}z}{z-\gamma_{n-1}}\phi_{n-1}^* \in \mathcal{L}_{n-2}.$$

This can be written as $N(z)/[(z-\gamma_{n-1})\pi_{n-1}(z)]$. Thus the c_n and d_n are defined by the conditions $N(\gamma_{n-1}) = N(1/\overline{\gamma}_{n-1}) = 0$. If we denote $\phi_k = p_k/\pi_k$ and thus $\phi_k^* = p_k^*\varsigma_k/\pi_k$, it is clear that

$$N(z) = p_n(z) - d_n(z-\gamma_{n-1})p_{n-1}(z) - c_n(1-\overline{\gamma}_{n-1}z)p_{n-1}^*(z)\varsigma_{n-1}.$$

Thus the first condition gives

$$c_n = \frac{\overline{\varsigma}_{n-1}p_n(\gamma_{n-1})}{(1-|\gamma_{n-1}|^2)p_{n-1}^*(\gamma_{n-1})}$$

and the second one

$$d_n = \frac{p_n(1/\overline{\gamma}_{n-1})}{(1/\overline{\gamma}_{n-1}-\gamma_{n-1})p_{n-1}(1/\overline{\gamma}_{n-1})} = \frac{\overline{p_n^*(\gamma_{n-1})}}{(1-|\gamma_{n-1}|^2)\overline{p_{n-1}^*(\gamma_{n-1})}}.$$

Note that $p_{n-1}^*(\gamma_{n-1})$ cannot be zero, and that also $p_n^*(\gamma_{n-1})$ does not vanish by our assumption of regularity.

Furthermore, it is not difficult to show that $\phi \perp \mathcal{L}_{n-2}$, so that it must be identically zero. Thus

$$\phi_n(z) = d_n\overline{\sigma}_{n-1}\frac{1-\overline{\gamma}_{n-1}z}{1-\overline{\gamma}_n z}[\zeta_{n-1}(z)\phi_{n-1}(z) + \lambda_n\phi_{n-1}^*(z)],$$

with

$$\lambda_n = \eta_n\frac{p_n(\gamma_{n-1})}{\overline{p_n^*(\gamma_{n-1})}}, \quad \eta_n = \overline{\varsigma}_{n-1}\sigma_{n-1}\frac{\overline{p_{n-1}^*(\gamma_{n-1})}}{p_{n-1}^*(\gamma_{n-1})}.$$

By taking the ($)^*$ transform (in $\mathcal{L}_n$) we obtain

$$\phi_n^*(z) = \overline{d}_n \sigma_n \frac{1-\overline{\gamma}_{n-1} z}{1-\overline{\gamma}_n z} [\overline{\lambda}_n \zeta_{n-1}(z)\phi_{n-1}(z) + \phi_{n-1}^*(z)].$$

This proves the recurrence by taking $e_n = |d_n|$ and $\overline{\eta}_{n1} = \sigma_{n-1}\mathfrak{u}(d_n)$.
It remains to show that the initial step for $n = 1$ is true. Since $\phi_0 = \phi_0^* = 1$, then in case $\gamma_0 = \alpha_0 = 0$, hence $\zeta_0 = z$, we have

$$\phi_1(z) = e_1\eta_{11}\frac{z+\lambda_1}{1-\overline{\gamma}_1 z} \quad \text{and} \quad \phi_1^*(z) = e_1\eta_{12}\frac{\overline{\lambda}_1 z + 1}{1-\overline{\gamma}_1 z}.$$

Thus

$$p_1(z) = e_1\eta_{11}(z+\lambda_1) \quad \text{and} \quad p_1^*(z) = e_1\overline{\eta}_{11}(\overline{\lambda}_1 z + 1).$$

This implies that λ_1 is indeed given by the general formula because

$$\lambda_1 = \eta_1 \frac{p_1(\gamma_0)}{\overline{p_1^*(\gamma_0)}} = \frac{p_1(0)}{\overline{p_1^*(0)}} = \frac{e_1\eta_{11}\lambda_1}{e_1\eta_{11}}.$$

In case $\gamma_0 = \beta_0 = \infty$, then $\zeta_0 = 1/z$, so that

$$\phi_1(z) = e_1\eta_{11}\frac{-1-\lambda_1 z}{1-\overline{\gamma}_1 z} \quad \text{and} \quad \phi_1^*(z) = e_1\eta_{12}\frac{-\overline{\lambda}_1 - z}{1-\overline{\gamma}_1 z},$$

so that

$$p_1(z) = -e_1\eta_{11}(1+\lambda_1 z) \quad \text{and} \quad p_1^*(z) = -e_1\overline{\eta}_{11}\overline{\sigma}_1(\overline{\lambda}_1 + z)$$

and again λ_1 is given by the general formula

$$\lambda_1 = \eta_1 \frac{p_1(\gamma_0)}{\overline{p_1^*(\gamma_0)}} = 1\frac{p_1(\infty)}{\overline{p_1^*(\infty)}} = \frac{e_1\eta_{11}\lambda_1}{e_1\eta_{11}}.$$

This proves the theorem. $\square$

Remark 5.2 If $\nu \in \{\alpha, \beta\}$ we could rewrite λ_n^ν in terms of ϕ_n^ν because by dividing and multiplying with the appropriate denominators π_n^ν one gets

$$\lambda_n^\nu = \eta_n^\nu \frac{\phi_n^\nu(\nu_{n-1})}{\overline{\phi_n^{\nu*}(\nu_{n-1})}}, \quad \eta_n^\nu = \sigma_{n-1}\overline{\sigma}_n \frac{(1-\overline{\nu}_n\nu_{n-1})}{(1-\nu_n\overline{\nu}_{n-1})} \frac{\overline{\phi_{n-1}^{\nu*}(\nu_{n-1})}}{\phi_{n-1}^{\nu*}(\nu_{n-1})}, \quad n \geq 1.$$

Note that also this $\eta_n^\nu \in \mathbb{T}$, but it differs from the η_n^ν in the previous theorem. However if $\nu = \gamma$, then this expression has the disadvantage that γ_{n-1} could be equal to $1/\overline{\gamma}_n$ or it could be equal to a pole of ϕ_n in which case it would not make sense to evaluate these expressions in γ_{n-1}. The latter expressions only make sense if we interpret them as limiting values

$$\frac{\phi_n^\nu(\nu_{n-1})}{\phi_n^{\nu*}(\nu_{n-1})} = \lim_{z\to\nu_{n-1}} \frac{\phi_n^\nu(z)}{\phi_n^{\nu*}(z)} \quad \text{and} \quad \frac{(1-\overline{\nu}_n\nu_{n-1})}{(1-\nu_n\overline{\nu}_{n-1})}\frac{\overline{\phi_{n-1}^{\nu*}(\nu_{n-1})}}{\phi_{n-1}^{\nu*}(\nu_{n-1})} = \lim_{z\to\nu_{n-1}} \frac{(1-\overline{\nu}_n z)}{(1-\nu_n\overline{z})}\frac{\overline{\phi_{n-1}^{\nu*}(z)}}{\phi_{n-1}^{\nu*}(z)},$$

where one has to assume that $\lim_{\xi\to 0}[\xi/\overline{\xi}] = 1$. We shall from now on occasionally use these expressions with this interpretation, but the expressions for λ_n^ν from Theorem 5.1 using the numerators are more reliable.

Note that λ_n^α is a Blaschke product with all its zeros in $\mathbb{D}$ that is evaluated at $\alpha_{n-1} \in \mathbb{D}$ and therefore $\lambda_n^\alpha \in \mathbb{D}$. Similarly, λ_n^β is a Blaschke product with all its zeros in $\mathbb{E}$ that is evaluated at $\beta_{n-1} \in \mathbb{E}$ so that also $\lambda_n^\beta \in \mathbb{D}$. Since the zeros of ϕ_n are the zeros of ϕ_n^α if $n \in \mathbb{a}_n$ and they are the zeros of ϕ_n^β if $n \in \mathbb{b}_n$, it follows that if n and $n-1$ are both in $\mathbb{a}_n$ or both in $\mathbb{b}_n$, then $\lambda_n \in \mathbb{D}$ but if $n \in \mathbb{a}_n$ and $n-1 \in \mathbb{b}_n$ or vice versa, then $\lambda_n \in \mathbb{E}$. This explains that

$$(e_n^\nu)^2 = \frac{1-|\nu_n|^2}{1-|\nu_{n-1}|^2}\frac{1}{1-|\lambda_n^\nu|^2} > 0. \tag{5.1}$$

and we can choose e_n as the positive square root of this expression. The above expression is derived in [5, Thm. 4.1.2] for the case $\nu = \alpha$ by using the CD relations. By symmetry, this also holds for $\nu = \beta$. For $\nu = \gamma$, the same relation can be obtained by stripping the denominators as we explained after the proof of the CD-relation in Section 4

What goes wrong with the recurrence relation when ϕ_n is not regular? From the proof of Theorem 5.1, it follows that then $d_n = 0$. We still have the relation

$$\phi_n(z) = \frac{\overline{\sigma}_{n-1}}{(1-|\gamma_{n-1}|^2)\overline{p_{n-1}^*(\gamma_{n-1})}}\left[p_n^*(\gamma_{n-1})\zeta_{n-1}(z)\phi_{n-1}(z) + s_{n-1}p_n(\gamma_{n-1})\phi_{n-1}^*(z)\right]$$

with $s_{n-1} = \frac{\varsigma_{n-1}p_{n-1}^*(\gamma_{n-1})}{\overline{\sigma}_{n-1}\overline{p_{n-1}^*(\gamma_{n-1})}} \in \mathbb{T}$ and $p_n^*(\gamma_{n-1}) = 0$. Thus, there is some positive constant e_n and some $\eta_{n1} \in \mathbb{T}$ such that

$$\phi_n(z) = e_n\eta_{n1}\frac{\varpi_{n-1}(z)}{\varpi_n(z)}\left[0\;\zeta_{n-1}(z)\phi_{n-1}(z) + \phi_{n-1}^*(z)\right].$$

I.e., the first term in the sum between square brackets vanishes. Applying Theorem 5.1 in this case would give $\lambda_n = \infty$, and the previous relations show that we only have to replace in Theorem 5.1 the matrix

$$\begin{bmatrix} 1 & \lambda_n^\nu \\ \overline{\lambda_n^\nu} & 1 \end{bmatrix} \quad \text{by} \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

This is conform with how we have dealt with ∞ so far where the rule of thumb was to set $a - \nu b = -b$ if $\nu = \infty$. So let us therefore also use Theorem 5.1 with this interpretation when ϕ_n is not regular and thus $\lambda_n = \infty$. With the expressions at the end of Section 4 it can also be shown that in this case

$$e_n^2 = -\frac{1 - |\gamma_n|^2}{1 - |\gamma_{n-1}|^2} > 0$$

Note that this corresponds to replacing $1 - |\lambda_n|^2$ when $\lambda_n = \infty$ by -1. Since this nonregular situation can only occur when $(1-|\gamma_n|)(1-|\gamma_{n-1}|) < 0$, this expression for e_n^2 is indeed positive. A similar rule can be applied if γ_n or γ_{n-1} is infinite, just replace in this or previous expression $1 - |\infty|^2$ by -1. The positivity of the expressions for e_n^2 also follows from the following result.

Theorem 5.3 *The Szegő parameters satisfy for all $n \geq 1$*
If $\gamma_n = \alpha_n$ and $\gamma_{n-1} = \alpha_{n-1}$ then $\lambda_n = \lambda_n^\alpha = \overline{\lambda_n^\beta} \in \mathbb{D}$.
If $\gamma_n = \beta_n$ and $\gamma_{n-1} = \beta_{n-1}$ then $\lambda_n = \lambda_n^\beta = \overline{\lambda_n^\alpha} \in \mathbb{D}$.
If $\gamma_n = \alpha_n$ and $\gamma_{n-1} = \beta_{n-1}$ then $\lambda_n = 1/\overline{\lambda_n^\beta} = 1/\lambda_n^\alpha \in \mathbb{E}$.
If $\gamma_n = \beta_n$ and $\gamma_{n-1} = \alpha_{n-1}$ then $\lambda_n = 1/\overline{\lambda_n^\alpha} = 1/\lambda_n^\beta \in \mathbb{E}$.

PROOF. Suppose $\gamma_n = \alpha_n$ and $\gamma_{n-1} = \alpha_{n-1}$, then by Theorems 5.1 and 3.8, or better still by Lemma 3.7,

$$\lambda_n = \left(\overline{\varsigma}_{n-2}\frac{\overline{p_{n-1}^*(\alpha_{n-1})}}{p_{n-1}^*(\alpha_{n-1})}\right)\frac{p_n(\alpha_{n-1})}{p_n^*(\alpha_{n-1})} = \left(\overline{\varsigma}_{n-2}\frac{\overline{p_{n-1}^{\alpha*}(\alpha_{n-1})}}{p_{n-1}^{\alpha*}(\alpha_{n-1})}\right)\frac{p_n^\alpha(\alpha_{n-1})}{p_n^{\alpha*}(\alpha_{n-1})} = \lambda_n^\alpha.$$

When using $p_n^\alpha(z) = \varsigma_n\, p_n^{\beta*}(z)\prod_{j=1}^n(-|\alpha_j|)$ and $\alpha_j = 1/\overline{\beta}_j$, the previous relation becomes

$$\begin{aligned}
\lambda_n &= \left(\overline{\varsigma}_{n-2}\frac{\varsigma_{n-1}\overline{p_{n-1}^\beta(1/\overline{\beta}_{n-1})}}{\overline{\varsigma}_{n-1}p_{n-1}^\beta(1/\overline{\beta}_{n-1})}\right)\frac{p_n^{\beta*}(1/\overline{\beta}_{n-1})}{p_n^\beta(1/\overline{\beta}_{n-1})}\\
&= \left(\sigma_{n-1}^2\varsigma_{n-2}\frac{p_n^{\beta*}(\beta_{n-1})}{\overline{p_n^{\beta*}(\beta_{n-1})}}\,\frac{\overline{\beta}_{n-1}^{n-1}}{\beta_{n-1}^{n-1}}\right)\frac{\overline{p_n^\beta(\beta_{n-1})}}{p_n^{\beta*}(\beta_{n-1})}\,\frac{\beta_{n-1}^n}{\overline{\beta}_{n-1}^n}\\
&= \sigma_{n-1}^2\frac{\beta_{n-1}}{\overline{\beta}_{n-1}}\overline{\lambda_n^\beta} = \overline{\lambda_n^\beta}.
\end{aligned}$$

The proof for $\gamma_n = \beta_n$ and $\gamma_{n-1} = \beta_{n-1}$, $n \geq 1$ is similar.

Next consider $\gamma_n = \alpha_n$ and $\gamma_{n-1} = \beta_{n-1}$, then

$$\lambda_n = \left(\overline{\varsigma}_{n-2} \frac{\overline{p_{n-1}^{\beta*}(\beta_{n-1})}}{p_{n-1}^{\beta*}(\beta_{n-1})} \right) \frac{p_n^\alpha(\beta_{n-1})}{p_n^{\alpha*}(\beta_{n-1})} = \eta_n^\beta \frac{p_n^{\beta*}(\beta_{n-1})}{p_n^\beta(\beta_{n-1})} = \frac{\eta_n^\beta \overline{\eta_n^\beta}}{\overline{\lambda_n^\beta}} = \frac{1}{\overline{\lambda_n^\beta}} = \frac{1}{\lambda_n^\alpha}.$$

The remaining case $\gamma_n = \beta_n$ and $\gamma_{n-1} = \alpha_{n-1}$, is again similar. $\square$

Another solution of the recurrence relation is formed by the functions of the second kind. Like in the classical case (i.e., for $\nu = \alpha$) we can introduce them for $\nu \in \{\alpha, \beta, \gamma\}$ by

$$\psi_n^\nu(z) = \int_{\mathbb{T}} [D(z,t)\phi_n^\nu(t) - E_n(z,t)\phi_n^\nu(z)] \, \mathrm{d}\mu(t).$$

where $D(t,z) = \frac{t+z}{t-z}$ and $E(t,z) = D(t,z) + 1 = \frac{2t}{t-z}$. This results in

$$\psi_0^\nu = 1 \quad \text{and} \quad \psi_n^\nu(z) = \int_{\mathbb{T}} D(t,z)[\phi_n^\nu(t) - \phi_n^\nu(z)] \, \mathrm{d}\mu(t), \quad n \geq 1$$

which may be generalized to

$$\psi_n^\nu(z) f(z) = \int_{\mathbb{T}} D(t,z)[\phi_n^\nu(t) f(t) - \phi_n^\nu(z) f(z)] \, \mathrm{d}\mu(t), \quad n \geq 1 \quad \text{with } f \text{ arbitrary in } \mathcal{L}_{(n-1)*}^\nu.$$

It also holds that

$$\psi_n^{\nu*}(z) g(z) = \int_{\mathbb{T}} D(t,z)[\phi_n^{\nu*}(t) g(t) - \phi_n^{\nu*}(z) g(z)] \, \mathrm{d}\mu(t), \quad n \geq 1 \quad \text{with } g \text{ arbitrary in } \mathcal{L}_{n*}^\nu(\nu_{n*}).$$

Recall that $\mathcal{L}_{n*}^\nu(\nu_{n*})$ is the space of all functions in $\mathcal{L}_{n*}^\nu$ that vanish for $z = \nu_{n*} = 1/\overline{\nu}_n$. This space is spanned by $\{B_k^\nu / B_n^\nu : k = 0, \ldots, n-1\}$ if $\nu \in \{\alpha, \beta\}$. For $\nu = \gamma$, the space is spanned by (see Lemma 3.1)

$$\mathcal{L}_{n*}(\gamma_{n*}) = \operatorname{span} \left\{ \frac{B_k}{\dot{B}_n^\alpha \dot{B}_n^\beta} \right\}_{k=0}^{n-1} = \operatorname{span} \left\{ \frac{B_k^\alpha}{\zeta_n \dot{B}_{n-1}^\alpha} \right\}_{k=0}^{n-1} = \operatorname{span} \left\{ \frac{B_k^\beta}{\zeta_n \dot{B}_{n-1}^\beta} \right\}_{k=0}^{n-1}.$$

Theorem 5.4 *The following relations for the functions of the second kind hold for $n \geq 0$.*

$$\psi_n = (\psi_n^\alpha) \dot{B}_n^\beta = (\psi_n^\beta)^* \dot{B}_n^\alpha \quad \text{and} \quad \psi_n^* = (\psi_n^\alpha)^* \dot{B}_n^\beta = (\psi_n^\beta) \dot{B}_n^\alpha \quad \text{if } \gamma_n = \alpha_n$$

while

$$\psi_n = (\psi_n^\beta) \dot{B}_n^\alpha = (\psi_n^\alpha)^* \dot{B}_n^\beta \quad \text{and} \quad \psi_n^* = (\psi_n^\beta)^* \dot{B}_n^\alpha = (\psi_n^\alpha) \dot{B}_n^\beta \quad \text{if } \gamma_n = \beta_n.$$

We assume the normalization of Theorem 3.8.

PROOF. This is trivial for $n = 0$, hence suppose $n \geq 1$ and $\gamma_n = \alpha_n$ then

$$\begin{aligned}\psi_n(z) &= \int_{\mathbb{T}} D(t,z)[\phi_n(t) - \phi_n(z)]\,\mathrm{d}\mu(t) = \int_{\mathbb{T}} D(t,z)[\phi_n^\alpha(t)\dot{B}_n^\beta(t) - \phi_n^\alpha(z)\dot{B}_n^\beta(z)]\,\mathrm{d}\mu(t)\\ &= \psi_n^\alpha(z)\dot{B}_n^\beta(z)\end{aligned}$$

because

$$\dot{B}_n^\beta(z) = \dot{B}_{n-1}^\beta(z) = \prod_{j\in\mathbb{b}_{n-1}} \zeta_j^\beta(z) = \prod_{j\in\mathbb{b}_{n-1}} \zeta_{j*}^\alpha(z) \in \mathcal{L}_{(n-1)*}^\alpha.$$

Moreover, using $\phi_{n*}^\alpha = \phi_n^\beta$ we also have

$$\begin{aligned}\psi_n(z) &= \int_{\mathbb{T}} D(t,z)[\phi_n(t) - \phi_n(z)]\,\mathrm{d}\mu(t) = \int_{\mathbb{T}} D(t,z)[\phi_n^\alpha(t)\dot{B}_n^\beta(t) - \phi_n^\alpha(z)\dot{B}_n^\beta(z)]\,\mathrm{d}\mu(t)\\ &= \int_{\mathbb{T}} D(t,z)[\phi_{n*}^\beta(t)\dot{B}_n^\beta(t) - \phi_{n*}^\beta(z)\dot{B}_n^\beta(z)]\,\mathrm{d}\mu(t)\\ &= \int_{\mathbb{T}} D(t,z)[\phi_n^{\beta*}(t)\dot{B}_n^\alpha(t) - \phi_n^{\beta*}(z)\dot{B}_n^\alpha(z)]\,\mathrm{d}\mu(t)\end{aligned}$$

and since $\dot{B}_n^\alpha \in \mathcal{L}_{n*}^\beta(\beta_{n*})$, we also get the second part: $\psi_n = (\psi_n^\beta)^*\dot{B}_n^\alpha$.

Moreover $\psi_n^* = [\psi_n^\alpha \dot{B}_n^\beta]_* \dot{B}_n^\alpha \dot{B}_n^\beta = \psi_{n*}^\alpha \dot{B}_n^\alpha \dot{B}_n^\beta \dot{B}_{n*}^\beta = [\psi_{n*}^\alpha \dot{B}_n^\alpha / \dot{B}_n^\beta]\dot{B}_n^\beta = \psi_n^{\alpha*}\dot{B}_n^\beta$. It follows in a similar way that $\psi_n^* = \psi_n^\beta \dot{B}_n^\alpha$.

The case $\gamma_n = \beta_n$ is proved similarly. $\square$

Corollary 5.5 *With the same notation as in the previous theorem, we can derive for the associated functions*

$$P_n^\nu(z,\tau_n) = \psi_n^\nu(z) - \tau_n\psi_n^{\nu*}(z), \quad \nu \in \{\alpha,\beta,\gamma\}$$

that for $n \geq 1$

$$P_n(z,\tau_n) = \dot{B}_n^\beta(z)P_n^\alpha(z,\tau_n) = -\tau_n\dot{B}_n^\alpha(z)P_n^\beta(z,\overline{\tau}_n) \quad \textit{if } \gamma_n = \alpha_n$$

while

$$P_n(z,\tau_n) = \dot{B}_n^\alpha(z)P_n^\beta(z,\tau_n) = -\tau_n\dot{B}_n^\beta(z)P_n^\alpha(z,\overline{\tau}_n) \quad \textit{if } \gamma_n = \beta_n.$$

These functions satisfy the recurrence relation

$$\begin{bmatrix} \psi_n^\nu(z) \\ -\psi_n^{\nu*}(z) \end{bmatrix} = N_n^\nu \frac{\varpi_{n-1}^\nu(z)}{\varpi_n^\nu(z)} \begin{bmatrix} 1 & \lambda_n^\nu \\ \overline{\lambda}_n^\nu & 1 \end{bmatrix} \begin{bmatrix} \zeta_{n-1}^\nu(z) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \psi_{n-1}^\nu(z) \\ -\psi_{n-1}^{\nu*}(z) \end{bmatrix}, \quad n \geq 1$$

with $\psi_0^\nu = \psi_0^{\nu *} = 1$ and all other quantities as in Theorem 5.1.

6 Para-orthogonal rational functions

The result of Corollary 3.10 is important to conclude that the zeros of the para-orthogonal rational functions defined as

$$Q_n^\nu(z, \tau_n^\nu) = \phi_n^\nu(z) + \tau_n^\nu \phi_n^{\nu *}(z), \quad \tau_n^\nu \in \mathbb{T}, \quad \nu \in \{\alpha, \beta, \gamma\}$$

will have similar properties (simple and on $\mathbb{T}$) no matter whether $\nu = \alpha$, β or γ.

Theorem 6.1 *The para-orthogonal rational function $Q_n(z) = Q_n(z, \tau_n)$, $\tau_n \in \mathbb{T}$ with $n \geq 2$ is orthogonal to $\dot{\zeta}_n^\alpha \mathcal{L}_{n-1} \cap \dot{\zeta}_n^\beta \mathcal{L}_{n-1} = \zeta_n \mathcal{L}_{n-1} \cap \mathcal{L}_{n-1} = \mathcal{L}_{n-1}(\gamma_n)$ with*

$$\mathcal{L}_{n-1}(\gamma_n) = \{f \in \mathcal{L}_{n-1} : f(\gamma_n) = 0\} = \left\{ \frac{\varpi_n^*(z) p_{n-2}(z)}{\pi_{n-1}(z)} : p_{n-2} \in \mathcal{P}_{n-2} \right\}.$$

Recall that $\varpi_n^*(z) = z - \gamma_n$ if $\gamma_n \neq \infty$, and $\varpi_n^*(z) = -1$ if $\gamma_n = \infty$.

PROOF. Suppose $\gamma_n = \alpha_n$ then $\dot{\zeta}_n^\alpha = \zeta_n^\alpha$ and $\dot{\zeta}_n^\beta = 1$. Hence $\phi_n \perp \mathcal{L}_{n-1}$ and $\phi_n^* \perp \zeta_n^\alpha \mathcal{L}_{n-1}$ and therefore $Q_n \perp \mathcal{L}_{n-1} \cap \zeta_n^\alpha \mathcal{L}_{n-1} = \mathcal{L}_{n-1}(\alpha_n)$.

Still assuming that $\gamma_n = \alpha_n$, then $\phi_n = \phi_n^\alpha \dot{B}_n^\beta$ and $\phi_n^* = (\phi_n^\alpha)^* \dot{B}_n^\beta$. Thus

$$Q_n(\cdot, \tau_n) = \dot{B}_n^\beta \left[\phi_n^\alpha + \tau_n (\phi_n^\alpha)^*\right] = \dot{B}_n^\beta Q_n^\alpha(\cdot, \tau_n).$$

So it is clear that the zeros of $Q_n(z, \tau_n)$ are the zeros of $Q_n^\alpha(z, \tau_n)$ which are all simple and on $\mathbb{T}$ (and hence will not cancel against any of the poles). In a similar way one has

$$Q_n(\cdot, \tau_n) = \phi_n^{\beta *} \dot{B}_n^\alpha + \tau_n \phi_n^\beta \dot{B}_n^\alpha = \tau_n \dot{B}_n^\alpha \left[\phi_n^\beta + \overline{\tau}_n \phi_n^{\beta *}\right] = \tau_n \dot{B}_n^\alpha Q_n^\beta(\cdot, \overline{\tau}_n).$$

The proof for $\gamma_n = \beta_n$ is similar. $\square$

As a direct consequence we get

Corollary 6.2 *Take $n \geq 1$, $\tau_n \in \mathbb{T}$ and $\nu \in \{\alpha, \beta, \gamma\}$ then the zeros of $Q_n^\nu(z, \tau_n)$ are all simple and on $\mathbb{T}$.*

In particular we have

$$Q_n(z, \tau_n) = \dot{B}_n^\beta(z) Q_n^\alpha(z, \tau_n) = \tau_n \dot{B}_n^\alpha(z) Q_n^\beta(z, \overline{\tau}_n), \quad \textit{if} \;\; \gamma_n = \alpha_n,$$

and

$$Q_n(z, \tau_n) = \dot{B}_n^\alpha(z) Q_n^\beta(z, \tau_n) = \tau_n \dot{B}_n^\beta(z) Q_n^\alpha(z, \overline{\tau}_n), \quad \textit{if} \;\; \gamma_n = \beta_n.$$

These PORF properties are important because the zeros will deliver the nodes for the rational Szegő quadrature as we will explain in the next section.

For further reference we give the following Property.

Proposition 6.3 *For $n \geq 1$ and $\nu \in \{\alpha, \beta, \gamma\}$, the PORF satisfy, using the notation of Theorem 5.1*

$$Q_n^\nu(z, \tau_n) = c_n^\nu \frac{\varpi_{n-1}^\nu(z)}{\varpi_n^\nu(z)} \left[\zeta_{n-1}^\nu(z) \phi_{n-1}^\nu(z) + \tilde{\tau}_n^\nu \phi_{n-1}^{\nu*}(z) \right]$$

with

$$c_n^\nu = e_n^\nu(\eta_{n1} + \eta_{n2}\tau_n \lambda_n^\nu), \quad \tilde{\tau}_n^\nu = \frac{\hat{\tau}_n + \lambda_n^\nu}{1 + \hat{\tau}_n \overline{\lambda}_n^\nu} \in \mathbb{T}, \quad \hat{\tau}_n = \overline{\eta}_{n1} \eta_{n2} \tau_n.$$

PROOF. Just take the recurrence relation of Theorem 5.1 and premultiply it with $[1 \;\; \tau_n]$. After re-arrangement of the terms, the expression above will result.

The importance of this property is that in fact up to a constant nonzero factor, we can compute $Q_n^\nu(z, \tau_n)$ by exactly the same recurrence that gives ϕ_n^ν in terms of ϕ_{n-1}^ν and $\phi_{n-1}^{\nu*}$, except that we have to replace λ_n^ν by a unimodular factor $\tilde{\tau}_n^\nu$. When we are interested in the zeros of $Q_n^\nu(z, \tau_n)$, then the constant that is up in front does not matter since it is nonzero.

7 Quadrature

We start by proving that the subspace of rational functions in which the quadrature formulas will be exact only depends on the points $\{\alpha_k : k = 0, \ldots, n-1\}$ no matter whether the points α_k are introduced as a pole α_k or $\alpha_{k*} = 1/\overline{\alpha}_k$ in the sequence $\underline{\gamma}_n$.

Lemma 7.1 $\mathcal{R}_n \coloneqq \mathcal{L}_n \cdot \mathcal{L}_{n*} = \mathcal{R}_n^\alpha \coloneqq \mathcal{L}_n^\alpha \cdot \mathcal{L}_{n*}^\alpha = \mathcal{R}_n^\beta \coloneqq \mathcal{L}_n^\beta \cdot \mathcal{L}_{n*}^\beta$

PROOF. The space $\mathcal{L}_n$ contains rational functions of degree n whose denominator has zeros in $\{\beta_j : j \in \mathbb{a}_n\} \cup \{\alpha_j : j \in \mathbb{b}_n\}$. The space $\mathcal{L}_{n*}$ contains rational functions of degree n whose denominator has zeros in $\{\beta_j : j \in \mathbb{b}_n\} \cup \{\alpha_j : j \in \mathbb{a}_n\}$. Thus the space $\mathcal{R}_n$ contains rational functions of degree $2n$ whose denominator has zeros in $\{\beta_j : j = 1, \ldots, n\} \cup \{\alpha_j : j = 1, \ldots, n\}$. Since the denominators of functions in $\mathcal{L}_n^\alpha$ have zeros in $\{\beta_j : j = 1, \ldots, n\}$ and functions in $\mathcal{L}_{n*}^\alpha$ have zeros in $\{\alpha_j : j = 1, \ldots, n\}$, the denominator of functions in $\mathcal{R}_n^\alpha$ have zeros in $\{\beta_j : j = 1, \ldots, n\} \cup \{\alpha_j : j = 1, \ldots, n\}$ so that $\mathcal{R}_n = \mathcal{R}_n^\alpha$. Of course a similar argument shows that also $\mathcal{R}_n = \mathcal{R}_n^\beta$. $\square$

As is well known from the classical case associated with the sequence $\underline{\boldsymbol{\alpha}}$, the rational Szegő quadrature formulas are of the form

$$I_n^\alpha(f) = \sum_{k=1}^{n} \omega_{nj}^\alpha f(\xi_{nj}^\alpha), \quad n \geq 1$$

with $\xi_{nj}^\alpha = \xi_{nj}^\alpha(\tau_n^\alpha)$ are the zeros of the para-orthogonal rational function $Q_n^\alpha(z, \tau_n^\alpha)$ with $\tau_n^\alpha \in \mathbb{T}$ and the weights are given by

$$\omega_{nj}^\alpha = \omega_{nj}^\alpha(\tau_n) = \frac{1}{\sum_{k=0}^{n-1} |\phi_k^\alpha(\xi_{nj}^\alpha(\tau_n^\alpha))|^2} > 0.$$

These formulas have maximal degree of exactness, meaning that

$$\int_{\mathbb{T}} f(t)\,\mathrm{d}\mu(t) = I_n^\alpha(f), \quad \forall f \in \mathcal{R}_{n-1}^\alpha$$

and that no n-point interpolatory quadrature formula with nodes on $\mathbb{T}$ and positive weights can be exact in a larger space of the form $\mathcal{L}_n^\alpha \cdot \mathcal{L}_{(n-1)*}^\alpha$ or $\mathcal{L}_{n*}^\alpha \cdot \mathcal{L}_{n-1}^\alpha$.

Our previous results show that taking an arbitrary order of selecting the $\underline{\boldsymbol{\gamma}}$ does not add new quadrature formulas. Indeed, if $\gamma_n = \alpha_n$ then the zeros of $Q_n(z, \tau_n)$, $Q_n^\alpha(z, \tau_n)$ and $Q_n^\beta(z, \overline{\tau}_n)$ are all the same, i.e., $\xi_{nj}(\tau_n) = \xi_{nj}^\alpha(\tau_n) = \xi_{nj}^\beta(\overline{\tau}_n)$, $j = 1, \ldots, n$. Dropping the dependency on τ from the notation, it is directly seen that

$$\sum_{k=0}^{n-1} |\phi_k(\xi_{nj})|^2 = \sum_{k=0}^{n-1} |\phi_k^\alpha(\xi_{nj}) \dot{B}_k^\beta(\xi_{nj})|^2 = \sum_{k=0}^{n-1} |\phi_k^\alpha(\xi_{nj})|^2 = \sum_{k=0}^{n-1} |\phi_k^\alpha(\xi_{nj}^\alpha)|^2$$

and thus we also have $\omega_{nj} = \omega_{nj}^{\alpha}$ and similarly $\omega_{nj} = \omega_{nj}^{\beta}$. Therefore $I_n = I_n^{\alpha} = I_n^{\beta}$ for appropriate choices of the defining parameters, i.e.,

$$\begin{aligned} \tau_n^{\alpha} &= \tau_n \quad \text{and} \quad \tau_n^{\beta} = \overline{\tau}_n \quad \text{if } \gamma_n = \alpha_n \\ \tau_n^{\beta} &= \tau_n \quad \text{and} \quad \tau_n^{\alpha} = \overline{\tau}_n \quad \text{if } \gamma_n = \beta_n. \end{aligned}$$

An alternative expression for the weights is also known and it will obviously also coincide for $\nu \in \{\alpha, \beta, \gamma\}$:

$$\omega_{nj}^{\nu} = \frac{1}{2\xi_{nj}^{\nu}} \left. \frac{P_n^{\nu}(z)}{\frac{\mathrm{d}}{\mathrm{d}z} Q_n^{\nu}(z)} \right|_{z=\xi_{nj}^{\nu}}$$

where Q_n^{ν} is the PORF with zeros $\{\xi_{nj}^{\nu}\}_{j=1}^{n}$ and P_n^{ν} the associated functions of the second kind as in Corollary 5.5. We have dropped again the obvious dependence on the parameter τ_n^{ν} from the notation.

A conclusion to be drawn from from this section is that whether we choose the sequence $\underline{\boldsymbol{\gamma}}$, $\underline{\boldsymbol{\alpha}}$ or $\underline{\boldsymbol{\beta}}$, the resulting quadrature formula we obtain is an n-point formula that is exact in the space $\mathcal{R}_{n-1} = \mathcal{L}_{n-1} \cdot \mathcal{L}_{(n-1)*}$. Such a quadrature formula is called an n-point rational Szegő quadrature and these are the only ones with the mentioned properties. They are unique up to the choice of the parameter $\tau_n \in \mathbb{T}$. Thus, to obtain the nodes and weights for a general sequence $\underline{\boldsymbol{\gamma}}$ and a choice for $\tau_n \in \mathbb{T}$, we can as well compute them for the sequence $\underline{\boldsymbol{\alpha}}$ and an appropriate $\tau_n^{\alpha} \in \mathbb{T}$ or for the sequence that alternates between one α and one β. The resulting quadrature will be the same. More about this later.

8 Block diagonal with Hessenberg blocks

The orthogonal polynomials are a special case of ORF obtained by choosing a sequence $\underline{\boldsymbol{\gamma}} = \underline{\boldsymbol{\alpha}}$ that contains only zeros. Another special case is given by the orthogonal Laurent polynomials (OLP) obtained by choosing an alternating sequence $\underline{\boldsymbol{\gamma}} = \{0, \infty, 0, \infty, 0, \infty, \ldots\}$. In [28], Velázquez described spectral methods for ORF on the unit circle that were based on these two special choices. The result was that the matrix representation of the shift operator $\mathcal{T}_{\mu} : L_{\mu}^2 \to L_{\mu}^2 : f(z) \mapsto zf(z)$ has a matrix representation that is a matrix Möbius transform of a structured matrix. The structure is a Hessenberg structure in the case of $\underline{\boldsymbol{\gamma}} = \underline{\boldsymbol{\alpha}}$ and it is a five-diagonal matrix (a so-called CMV matrix) for the sequence $\underline{\boldsymbol{\gamma}} = \{0, \beta_1, \alpha_2, \beta_3, \alpha_4, \ldots\}$. In the case of orthogonal (Laurent) polynomials the Möbius transform turns out to be just the identity and we get the plain Hessenberg matrix for the polynomials and the plain CMV matrix for the Laurent polynomials.

In [8], the OLP case was discussed using an alternative approach when the $0, \infty$ choice did not alternate nicely, but the order in which they were added was arbitrary. Then the structured matrix generalized to a so-called snake-shape matrix. This is in fact a generalization of both the Hessenberg and the CMV structures mentioned above. It is a block diagonal where the blocks alternate between upper and lower Hessenberg structure.

To illustrate this for our ORF, we start by using the approach of Velázquez in [28] to obtain this structure. In the next sections we shall use the linear algebra approach of [8] to analyze the operator aspects and the computational aspects for the quadrature formulas.

We start from the recurrence for the ϕ_k^α and transform it into a recurrence relation for the ϕ_k, which will eventually result in a representation of the shift operator.

To avoid a tedious book-keeping of normalizing constants, we will just exploit the fact that there are some constants such that certain dependencies hold. For example the recurrence

$$\phi_n^\alpha(z) = e_n \eta_{n1} \frac{\varpi_{n-1}^\alpha(z)}{\varpi_n^\alpha(z)} \left[\zeta_{n-1}^\alpha(z)\phi_{n-1}^\alpha(z) + \lambda_n^\alpha \phi_{n-1}^{\alpha*}(z)\right]$$

or equivalently

$$\phi_n^\alpha(z) = e_n \eta_{n1} \left[\sigma_{n-1} \frac{\varpi_{n-1}^{\alpha*}(z)}{\varpi_n^\alpha(z)} \phi_{n-1}^\alpha(z) + \lambda_n^\alpha \frac{\varpi_{n-1}^\alpha(z)}{\varpi_n^\alpha(z)} \phi_{n-1}^{\alpha*}(z)\right]$$

or

$$\varpi_{n-1}^{\alpha*}(z)\phi_{n-1}^\alpha(z) = e_n^{-1}\overline{\eta}_{n1}\overline{\sigma}_{n-1}\varpi_n^\alpha(z)\phi_n^\alpha(z) - \lambda_n^\alpha \overline{\sigma}_{n-1}\varpi_{n-1}^\alpha(z)\phi_{n-1}^{\alpha*}(z) \tag{8.1}$$

will be expressed as

$$\varpi_{n-1}^{\alpha*}\phi_{n-1}^\alpha \in \operatorname{span}\{\varpi_n^\alpha\phi_n^\alpha, \varpi_{n-1}^\alpha\phi_{n-1}^{\alpha*}\}. \tag{8.2}$$

Similarly, by combining the recurrence for ϕ_n^α and $\phi_n^{\alpha*}$ from Theorem 5.1 and eliminating ϕ_{n-1}^α, we have

$$\varpi_n^\alpha\phi_n^{\alpha*} \in \operatorname{span}\{\varpi_n^\alpha\phi_n^\alpha, \varpi_{n-1}^\alpha\phi_{n-1}^{\alpha*}\}. \tag{8.3}$$

Note that this relation always holds whether or not $\lambda_n^\alpha \neq 0$.

Now suppose that our sequence $\underline{\boldsymbol{\gamma}}$ has the following configuration

$$\underline{\boldsymbol{\gamma}} = \ldots, \beta_{k-1}, \alpha_k, \ldots, \alpha_n, \ldots, \alpha_{m-1}, \beta_m, \ldots \beta_{l-1}, \alpha_l, \ldots$$

with $2 \le k \le n < m < l$.

By using (8.2) and then repeatedly (8.3) we get

$$\varpi_n^{\alpha*}\phi_n^\alpha \in \operatorname{span}\{\varpi_{n+1}^\alpha\phi_{n+1}^\alpha, \ldots, \varpi_k^\alpha\phi_k^\alpha, \varpi_{k-1}^\alpha\phi_{k-1}^{\alpha*}\}.$$

Multiply this with $\dot{B}_n^\beta = \dot{B}_{n-1}^\beta = \cdots = \dot{B}_{k-1}^\beta$

$$\varpi_n^{\alpha*} \dot{B}_n^\beta \phi_n^\alpha \in \operatorname{span}\{\varpi_{n+1}^\alpha \dot{B}_n^\beta \phi_{n+1}^\alpha, \ldots, \varpi_k^\alpha \dot{B}_k^\beta \phi_k^\alpha, \varpi_{k-1}^\alpha \dot{B}_{k-1}^\beta \phi_{k-1}^{\alpha*}\},$$

and since by Theorem 3.8 $\phi_p = \dot{B}_p^\beta \phi_p^\alpha$ for $p = n, n-1, \ldots, k$ and $\phi_{k-1} = \dot{B}_{k-1}^\beta \phi_{k-1}^{\alpha*}$, this becomes

$$\varpi_n^{\alpha*} \phi_n \in \operatorname{span}\{\varpi_{n+1}^\alpha \dot{B}_n^\beta \phi_{n+1}^\alpha, \varpi_n^\alpha \phi_n, \ldots, \varpi_k^\alpha \phi_k, \varpi_{k-1}^\alpha \phi_{k-1}\}. \tag{8.4}$$

Thus if $n+1 < m$, then $\gamma_{n+1} = \alpha_{n+1}$, hence $\dot{B}_{n+1}^\beta = \dot{B}_n^\beta$ and $\phi_{n+1} = \dot{B}_{n+1}^\beta \phi_{n+1}^\alpha$, so that

$$\varpi_n^{\alpha*} \phi_n \in \operatorname{span}\{\varpi_p^\alpha \phi_p\}_{p=k-1}^{n+1}, \quad n+1 < m. \tag{8.5}$$

If $n+1 = m$, then $\gamma_{n+1} = \beta_m$ and we need to deal with the subsequence of β's in

$$\underline{\gamma} = \ldots, \alpha_{m-1}, \beta_m, \ldots, \beta_{l-1}, \alpha_l, \ldots$$

Therefore we note that

$$\varpi_m^\alpha = (z - \alpha_m) \frac{1 - \overline{\alpha}_m z}{z - \alpha_m} = (z - \alpha_m) \frac{z - \beta_m}{1 - \overline{\beta}_m z} \sigma_m^2 = \sigma_m \varpi_m^{\alpha*}(z) \zeta_m^\beta(z). \tag{8.6}$$

Thus

$$\varpi_m^\alpha \dot{B}_{m-1}^\beta \phi_m^\alpha = \sigma_m \varpi_m^{\alpha*} \dot{B}_m^\beta \phi_m^\alpha.$$

Using (8.2) again repeatedly we then see that

$$\begin{aligned}
\varpi_m^\alpha \dot{B}_{m-1}^\beta \phi_m^\alpha \in \dot{B}_m^\beta \operatorname{span}\{\varpi_m^\alpha \phi_m^{\alpha*}, \varpi_{m+1}^\alpha \phi_{m+1}^\alpha\} &= \dot{B}_m^\beta \operatorname{span}\{\varpi_m^\alpha \phi_m^{\alpha*}, \varpi_{m+1}^{\alpha*} \zeta_{m+1}^\beta \phi_{m+1}^\alpha\} \\
&= \operatorname{span}\{\varpi_m^\alpha \dot{B}_m^\beta \phi_m^{\alpha*}, \varpi_{m+1}^{\alpha*} \dot{B}_{m+1}^\beta \phi_{m+1}^\alpha\} \\
&= \cdots \\
&= \operatorname{span}\{\varpi_m^\alpha \dot{B}_m^\beta \phi_m^{\alpha*}, \ldots, \varpi_{l-1}^\alpha \dot{B}_{l-1}^\beta \phi_{l-1}^{\alpha*}, \varpi_l^\alpha \dot{B}_l^\beta \phi_l^\alpha\} \\
&= \operatorname{span}\{\varpi_m^\alpha \phi_m, \ldots, \varpi_{l-1}^\alpha \phi_{l-1}, \varpi_l^\alpha \phi_l\}
\end{aligned}$$

so that by plugging this into (8.4) with $n+1 = m$ gives

$$\varpi_n^{\alpha*} \phi_n \in \operatorname{span}\{\varpi_p^\alpha \phi_p\}_{p=k-1}^{l}, \quad n+1 = m. \tag{8.7}$$

The two relations (8.4) and (8.7) tell us how we should express $\varpi_n^{\alpha*} \phi_n$ in terms of the $\varpi_k^\alpha \phi_k$, $k \le n+1$ in the case that $\gamma_n = \alpha_n$.

The next step is to express $\varpi_n^{\alpha*} \phi_n$ in terms of $\{\varpi_p^\alpha \phi_p\}$, $p \ge n-1$ when $\gamma_n = \beta_n$. This goes along the same lines. Let us consider the mirror situation

$$\underline{\gamma} = \ldots, \beta_{l-1}, \alpha_l, \ldots, \alpha_{k-1}, \beta_k, \ldots, \beta_n, \ldots, \beta_{m-1}, \alpha_m, \ldots$$

with $2 \leq k \leq n < m < l$.

We first use (8.6) for $m = n$ to write

$$\varpi_n^{\alpha*}\phi_n = \varpi_n^{\alpha*}\dot{B}_n^{\beta}\phi_n^{\alpha*} = \overline{\sigma}_n\varpi_n^{\alpha}\dot{B}_{n-1}^{\beta}\phi_n^{\alpha*}$$

which implies that we shall need expressions for $\varpi_n^{\alpha}\phi_n^{\alpha*}$. We use repeatedly (8.3) in combination with the previous relation to get

$$\begin{aligned}
\varpi_n^{\alpha*}\phi_n \in \dot{B}_{n-1}^{\beta}\,&\mathrm{span}\{\varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha*}, \varpi_n^{\alpha}\phi_n^{\alpha}\}\\
&= \mathrm{span}\{\dot{B}_{n-1}^{\beta}\varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha*}, \dot{B}_n^{\beta}\varpi_n^{\alpha*}\phi_n^{\alpha}\}\\
&= \mathrm{span}\{\dot{B}_{n-1}^{\beta}\varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha*}, \dot{B}_n^{\beta}\varpi_n^{\alpha}\phi_n^{\alpha*}, \dot{B}_n^{\beta}\varpi_{n+1}^{\alpha}\phi_{n+1}^{\alpha}\}\\
&= \mathrm{span}\{\dot{B}_{n-1}^{\beta}\varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha*}, \dot{B}_n^{\beta}\varpi_n^{\alpha}\phi_n^{\alpha*}, \dot{B}_{n+1}^{\beta}\varpi_{n+1}^{\alpha*}\phi_{n+1}^{\alpha}\}\\
&= \mathrm{span}\{\dot{B}_{n-1}^{\beta}\varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha*}, \dot{B}_n^{\beta}\varpi_n^{\alpha}\phi_n^{\alpha*}, \dot{B}_{n+1}^{\beta}\varpi_{n+1}^{\alpha}\phi_{n+1}^{\alpha*}, \dot{B}_{n+1}^{\beta}\varpi_{n+2}^{\alpha}\phi_{n+2}^{\alpha}\}\\
&= \cdots\\
&= \mathrm{span}\{\dot{B}_{n-1}^{\beta}\varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha*}, \ldots, \dot{B}_{m-1}^{\beta}\varpi_{m-1}^{\alpha}\phi_{m-1}^{\alpha*}, \dot{B}_{m-1}^{\beta}\varpi_m^{\alpha}\phi_m^{\alpha}\}\\
&= \mathrm{span}\{\dot{B}_{n-1}^{\beta}\varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha*}, \varpi_n^{\alpha}\phi_n, \ldots, \varpi_{m-1}^{\alpha}\phi_{m-1}, \varpi_m^{\alpha}\phi_m\}
\end{aligned}$$

where in the last step we used $\gamma_m = \alpha_m$ and hence $\dot{B}_{m-1}^{\beta} = \dot{B}_m^{\beta}$ and $\dot{B}_{m-1}^{\beta}\phi_m^{\alpha} = \phi_m$. Thus if $n > k$, also $\gamma_{n-1} = \beta_{k-1}$ so that also the first term is $\varpi_{n-1}^{\alpha}\phi_{n-1}$ and we have found that

$$\varpi_n^{\alpha*}\phi_n \in \mathrm{span}\{\varpi_p^{\alpha}\phi_p^{\alpha}\}_{p=n-1}^{m}, \quad n > k. \tag{8.8}$$

Now we are left with the remaining case $n = k$, i.e., $\gamma_{n-1} = \alpha_{k-1}$, and so we are dealing with the α's in

$$\ldots, \beta_{l-1}\alpha_l, \ldots, \alpha_{m-1}, \beta_m \ldots$$

Using (8.3) repeatedly we get

$$\begin{aligned}
\varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha*} &\in \mathrm{span}\{\varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha}, \varpi_{n-2}^{\alpha}\phi_{n-2}^{\alpha*}\}\\
&= \mathrm{span}\{\varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha}, \varpi_{n-2}^{\alpha}\phi_{n-2}^{\alpha}, \varpi_{n-2}^{\alpha}\phi_{n-2}^{\alpha*}\}\\
&= \cdots\\
&= \mathrm{span}\{\varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha}, \varpi_{n-2}^{\alpha}\phi_{n-2}^{\alpha}, \ldots, \varpi_l^{\alpha}\phi_l^{\alpha}, \varpi_{l-1}^{\alpha}\phi_{l-1}^{\alpha*}\}.
\end{aligned}$$

After multiplying with $\dot{B}_l^{\beta} = \dot{B}_{l+1}^{\beta} = \cdots = \dot{B}_{n-1}^{\beta}$ we get

$$\varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha*} \in \mathrm{span}\{\varpi_p^{\alpha}\phi_p\}_{p=l-1}^{n-1}.$$

Plug this in our previous expression for $\varpi_n^{\alpha *}\phi_n$ and we arrive at

$$\varpi_n^{\alpha *}\phi_n \in \mathrm{span}\{\varpi_p^{\alpha}\phi_p\}_{p=l-1}^{m}, \quad n=k. \tag{8.9}$$

To summarize, (8.4) and (8.7) in the case $\gamma_n = \alpha_n$ and (8.8) and (8.9) in the case $\gamma_n = \beta_n$ show short recurrences for the ϕ_p that fully rely on the recursion for the α-related quantities: ϕ_n^{α} and $\phi_n^{\alpha *}$. They use only factors ϖ_p^{α} and $\varpi_p^{\alpha *}$ and in the relations, only the rational Szegő parameters λ_p^{α} are used.

This analysis should illustrate that as long as we are in a succession of α's we are building up an upper Hessenberg block. At the instant the α sequence switches to a β sequence one starts building a lower Hessenberg block, which switches back to upper Hessenberg when again α's enter the γ sequence etc. See Figure 1. Of course if there are only α's in the sequence, we end up with just an upper Hessenberg as in the classical case. If we alternate between one α and one β's we get the so called CMV type matrix which is a five-diagonal matrix with a characteristic block structure a given in Figure 2.

Suppose we start with a set of α's, then if we set $\mathcal{T} = z\mathcal{I}$ and

$$\mathcal{A} = \mathrm{diag}(\alpha_0, \alpha_1, \ldots)$$

and $\varpi_{\mathcal{A}}^{*} = \mathcal{T} - \mathcal{A}$ and $\varpi_{\mathcal{A}} = \mathcal{I} - \mathcal{T}\mathcal{A}^{*}$, then (at least formally) we have

$$[\phi_0, \phi_1, \ldots]\varpi_{\mathcal{A}}^{*} = [\phi_0, \phi_1, \ldots]\varpi_{\mathcal{A}}\hat{\mathcal{G}} \tag{8.10}$$

with $\hat{\mathcal{G}}$ having the structure shown in Figure 1. In the special case that the α's and β's alternate as in

$$\underline{\gamma} = \{\alpha_1, \beta_2, \alpha_3, \beta_4, \alpha_5, \ldots\},$$

we get the five-diagonal matrix as in [6] for polynomials and for ORF in [28]. Since for $k \geq 1$ each α_{2k-1}-column is the last in an α-block and each β_{2k}-column is the first in a β-block, we get a succession of 4×2 blocks that shift down by two rows as illustrated by Figure 2.

The particular role of the last column in an α-block and the first column in a β-block is due to the fact that we have chosen to derive these relations for the ϕ_k from the recurrence for the ϕ_k^{α} leading to factors $\varpi_{\mathcal{A}}$ and $\varpi_{\mathcal{A}}^{*}$ and the use of parameters ω_k^{α}. A symmetric derivation could have been given by starting from the ϕ_k^{β} recursion, in which case the β blocks will correspond to upper Hessenberg blocks and the α blocks to lower Hessenberg blocks. Then the longer recurrence would occur in the last column of the β-block and the first column of the α-block.

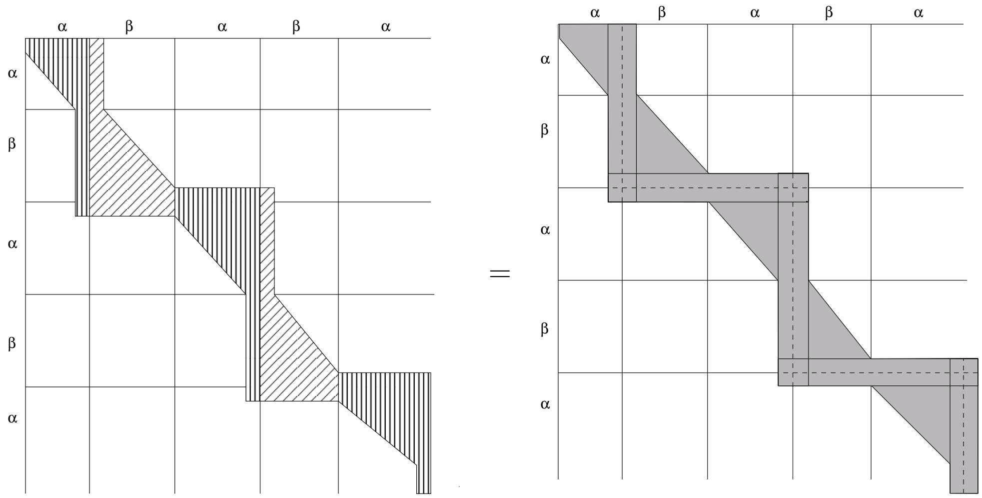

Fig. 1. Structure of the matrix $\hat{\mathcal{G}}$ which is the same as the structure of the matrix $\mathcal{G}_{\mathcal{A}}$.

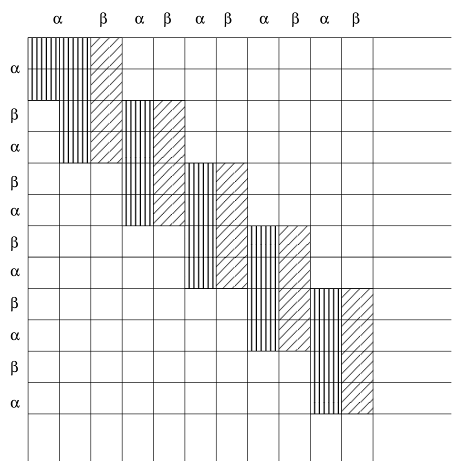

Fig. 2. Structure of the matrix $\hat{\mathcal{G}}_m$ when α's and β's alternate, which is the five-diagonal CMV matrix given in [28].

We can define as in [28] an operator Möbius transform

$$\tilde{\zeta}_{\mathcal{A}}(\mathcal{T}) = n_{\mathcal{A}}^{-1}\tilde{\varpi}^*_{\mathcal{A}}(\mathcal{T})\tilde{\varpi}_{\mathcal{A}}(\mathcal{T})^{-1}n_{\mathcal{A}^*}, \quad \begin{cases} \tilde{\varpi}_{\mathcal{A}}(\mathcal{T}) = \mathcal{I} + \mathcal{A}^*\mathcal{T} \\ \tilde{\varpi}^*_{\mathcal{A}}(\mathcal{T}) = \mathcal{T} + \mathcal{A} \\ n_{\mathcal{A}} = \varpi_{\mathcal{A}^*}(\mathcal{A}^*)^{1/2} = \sqrt{\mathcal{I} - \mathcal{A}\mathcal{A}^*} = \sqrt{\mathcal{I} - \mathcal{A}^*\mathcal{A}} = n_{\mathcal{A}^*} \\ \qquad\qquad = \operatorname{diag}(1, \sqrt{1-|\alpha_1|^2}, \sqrt{1-|\alpha_2|^2}, \ldots) \end{cases}$$

Then it is not difficult to see that (8.10) is equivalent with

$$[\phi_0(z), \phi_1(z), \phi_2(z), \ldots](z\mathcal{I} - \tilde{\zeta}_{\mathcal{A}}(\mathcal{G})) = 0, \quad \hat{\mathcal{G}} = n_{\mathcal{A}}^{-1}\mathcal{G}\, n_{\mathcal{A}^*} = \mathcal{G}_{\mathcal{A}}.$$

The matrices $\mathcal{G}$ and $\hat{\mathcal{G}}$ have the same form, but $\mathcal{G}$ is rescaled to be isometric. So we see intuitively that on $\mathcal{L} = \text{span}\{\phi_0, \phi_1, \ldots\}$ the operator $z\mathcal{I}$ should have a matrix representation $\tilde{\zeta}_{\mathcal{A}}(\mathcal{G})$ with respect to the ϕ-basis. We shall have a more careful discussion of this fact in the next section.

In this section have used the α-recursion to derive the form of the matrix. In principle, this could also be obtained from the β-recursion, but while it is rather simple to deal with $\alpha_j = 0$, it is much more tricky to deal with the corresponding $\beta_j = \infty$. Therefore we do not include that here.

It should also be clear that the alternating case (i.e., the case of the CMV representation) gives the smallest possible bandwidth.

9 Factorization of a general CMV matrix

It is well known (see [28,2] in the rational case and [8] for the polynomial case) that the Hessenberg matrix $\mathcal{G} = \mathcal{H}$ (obtained when $\underline{\boldsymbol{\gamma}} = \underline{\boldsymbol{\alpha}}$) can be written as an infinite product of Givens matrices, i.e.,

$$\mathcal{G} = \mathcal{H} = G_1 G_2 G_3 G_4 \cdots$$

where (the basis functions are orthogonal but we make abstraction of the finer normalization used)

$$G_k := \begin{bmatrix} I_{k-1} & 0 & 0 \\ 0 & \hat{G}_k & 0 \\ 0 & 0 & I_\infty \end{bmatrix}, \quad \tilde{G}_k := \begin{bmatrix} -\delta_k & \eta_k \\ \eta_k & \overline{\delta_k} \end{bmatrix}, \quad \forall k \geq 1,$$

with I_{k-1}, I_∞ are the identity matrices of sizes $(k-1)$ and ∞, respectively, I_0 is the empty matrix, $\delta_k = \lambda_k^\alpha$ is the k-th Szegő parameter and $\eta_k := \sqrt{1 - |\delta_k|^2}$.

Also the CMV matrix $\mathcal{C}^\varepsilon$ associated with the alternating sequence $\underline{\boldsymbol{\gamma}} = \underline{\boldsymbol{\varepsilon}}$ where $\gamma_{2k-1} = \alpha_{2k-1}$ and $\gamma_{2k} = \beta_{2k}$, for all $k \geq 1$ uses *the same* Givens transforms, but they are multiplied in a different order:

$$\mathcal{G} = \mathcal{C}_o\mathcal{C}_e = (\cdots G_9 G_7 G_5 G_3 G_1)\cdot(G_2 G_4 G_6 G_8 G_{10} \cdots) = (G_1 G_3 G_5 G_7 G_9 \cdots)\cdot(G_2 G_4 G_6 G_8 G_{10} \cdots).$$

To find explicit expressions for these elementary factors in our case, taking into account the proper normalization, requires a more detailed analysis. We start with the following

Theorem 9.1 *In the case $\underline{\boldsymbol{\gamma}} = \underline{\boldsymbol{\alpha}}$, we may rewrite the recurrence relation in this form*

$$\left[\frac{\varpi_{n-1}^{\alpha*}\phi_{n-1}^{\alpha}}{\sqrt{1-|\alpha_{n-1}|^2}} \middle| \frac{\varpi_n^{\alpha}\phi_n^{\alpha*}}{\sqrt{1-|\alpha_n|^2}} \right] = \left[\frac{\varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha*}}{\sqrt{1-|\alpha_{n-1}|^2}} \middle| \frac{\varpi_n^{\alpha}\phi_n^{\alpha}}{\sqrt{1-|\alpha_n|^2}} \right] \tilde{G}_n^{\alpha}$$

with

$$\tilde{G}_n^{\alpha} = \overline{\sigma}_{n-1}\overline{\eta}_{n1}^{\alpha} \begin{bmatrix} -\lambda_n^{\alpha}\eta_{n1}^{\alpha} & \sqrt{1-|\lambda_n^{\alpha}|^2} \\ \sqrt{1-|\lambda_n^{\alpha}|^2} & \overline{\lambda_n^{\alpha}}\overline{\eta}_{n1}^{\alpha} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sigma_n \end{bmatrix}$$

or

$$\left[\varpi_{n-1}^{\alpha*}\phi_{n-1}^{\alpha}|\varpi_n^{\alpha}\phi_n^{\alpha*}\right] = \left[\varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha*}|\varpi_n^{\alpha}\phi_n^{\alpha}\right]\hat{G}_n^{\alpha}$$

with

$$\hat{G}_n^{\alpha} = \begin{bmatrix} (1-|\alpha_{n-1}|^2)^{-1/2} & 0 \\ 0 & (1-|\alpha_n|^2)^{-1/2} \end{bmatrix} \tilde{G}_n^{\alpha} \begin{bmatrix} (1-|\alpha_{n-1}|^2)^{1/2} & 0 \\ 0 & (1-|\alpha_n|^2)^{1/2} \end{bmatrix}$$

PROOF. This is in fact an explicit form of what in the previous section was expressed as (8.2) and (8.3).

Taking the first line of the recurrence relation and solving for $\varpi_{n-1}^{\alpha*}\phi_{n-1}^{\alpha}$ we get (8.1), i.e.,

$$\varpi_{n-1}^{\alpha*}\phi_{n-1}^{\alpha} = (e_n^{\alpha})^{-1}\overline{\eta}_{n1}^{\alpha}\overline{\sigma}_{n-1}\varpi_n^{\alpha}\phi_n^{\alpha} - \lambda_n^{\alpha}\overline{\sigma}_{n-1}\varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha*}$$

Using $e_n^{\alpha} = \frac{\sqrt{1-|\alpha_n|^2}}{\sqrt{1-|\alpha_{n-1}|^2}}\frac{1}{\sqrt{1-|\lambda_n^{\alpha}|^2}}$ and $\eta_{n2}^{\alpha} = \overline{\eta}_{n1}^{\alpha}\overline{\sigma}_{n-1}\sigma_n$ this becomes

$$\frac{\varpi_{n-1}^{\alpha*}\phi_{n-1}^{\alpha}}{\sqrt{1-|\alpha_{n-1}|^2}} = \overline{\eta}_{n1}^{\alpha}\overline{\sigma}_{n-1}\left[-\lambda_n^{\alpha}\eta_{n1}^{\alpha}\frac{\varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha*}}{\sqrt{1-|\alpha_{n-1}|^2}} + \sqrt{1-|\lambda_n^{\alpha}|^2}\frac{\varpi_n^{\alpha}\phi_n^{\alpha}}{\sqrt{1-|\alpha_n|^2}}\right] \tag{9.1}$$

The second line of the recurrence relation is

$$\varpi_n^{\alpha}\phi_n^{\alpha*} = e_n^{\alpha}\eta_{n2}^{\alpha}\sigma_{n-1}\overline{\lambda_n^{\alpha}}\varpi_{n-1}^{\alpha*}\phi_{n-1}^{\alpha} + e_n^{\alpha}\eta_{n2}^{\alpha}\varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha*}.$$

Eliminate the term $\varpi_{n-1}^{\alpha*}\phi_{n-1}^{\alpha}$ with the first line of the recurrence and after inserting the value of e_n^α and rearranging, we get

$$\frac{\varpi_n^\alpha \phi_n^\alpha}{\sqrt{1-|\alpha_n|^2}} = \overline{\eta}_{n1}^\alpha \overline{\sigma}_{n-1}\left[\frac{\varpi_{n-1}^\alpha \phi_{n-1}^\alpha}{\sqrt{1-|\alpha_{n-1}|^2}}\sqrt{1-|\lambda_n^\alpha|^2}\sigma_n + \frac{\varpi_n^\alpha \phi_n^\alpha}{\sqrt{1-|\alpha_n|^2}}\overline{\lambda_n^\alpha}\overline{\eta}_{n1}^\alpha \sigma_n\right].$$

This is equivalent with the formula using $\tilde{G}_n^\alpha$ of the theorem. Derivation of the form $\hat{G}_n^\alpha$ is obvious. $\square$

Note that $\tilde{G}_n^\alpha$ is unitary but $\hat{G}_n^\alpha$ is not unless $|\lambda_n^\alpha| = 1$ or $\alpha_n = \alpha_{n-1}$.

Now suppose that $n \in \mathrm{a}_n$, i.e., $\gamma_n = \alpha_n$ then $\dot{B}_n^\beta = \dot{B}_{n-1}^\beta$ and hence

$$\left[\varpi_{n-1}^{\alpha*}\phi_{n-1}^\alpha \dot{B}_{n-1}^\beta \middle| \varpi_n^\alpha \phi_n^{\alpha*} \dot{B}_n^\beta\right] = \left[\varpi_{n-1}^\alpha \phi_{n-1}^{\alpha*} \dot{B}_{n-1}^\beta \middle| \varpi_n^\alpha \phi_n^\alpha \dot{B}_n^\beta\right] \hat{G}_n^\alpha. \tag{9.2}$$

If $n \in \mathbb{b}_n$, i.e., $\gamma_n = \beta_n$, then $\dot{B}_n^\beta = \dot{B}_{n-1}^\beta \zeta_n^\beta = \overline{\sigma}_n \dot{B}_{n-1}^\beta \frac{\varpi_n^\alpha}{\varpi_n^{\alpha*}}$.
Define then

$$\tilde{G}_n^\beta = \begin{bmatrix} \overline{\sigma}_n & 0 \\ 0 & 1 \end{bmatrix} \tilde{G}_n^\alpha \begin{bmatrix} \sigma_n & 0 \\ 0 & 1 \end{bmatrix} = \tilde{S}_n^* \tilde{G}_n^\alpha \tilde{S}_n, \quad \tilde{S}_n = \mathrm{diag}(\sigma_n, 1), \tag{9.3}$$

and as for $\hat{G}_n^\alpha$

$$\hat{G}_n^\beta = \begin{bmatrix} (1-|\alpha_{n-1}|^2)^{-1/2} & 0 \\ 0 & (1-|\alpha_n|^2)^{-1/2} \end{bmatrix} \tilde{G}_n^\beta \begin{bmatrix} (1-|\alpha_{n-1}|^2)^{1/2} & 0 \\ 0 & (1-|\alpha_n|^2)^{1/2} \end{bmatrix}$$

$$= \begin{bmatrix} \overline{\sigma}_n & 0 \\ 0 & 1 \end{bmatrix} \hat{G}_n^\alpha \begin{bmatrix} \sigma_n & 0 \\ 0 & 1 \end{bmatrix}.$$

so that

$$
\begin{aligned}
&\left[\,\varpi^{\alpha}_{n-1}\phi^{\alpha *}_{n-1}\dot{B}^{\beta}_{n-1}\middle|\varpi^{\alpha *}_{n}\phi^{\alpha}_{n}B^{\beta}_{n}\right]\hat{G}^{\beta}_{n}\\
&=\left[\,\varpi^{\alpha}_{n-1}\phi^{\alpha *}_{n-1}\dot{B}^{\beta}_{n-1}\middle|\varpi^{\alpha}_{n}\phi^{\alpha}_{n}\dot{B}^{\beta}_{n-1}\right]\overline{\sigma}_n\hat{G}^{\alpha}_{n}\begin{bmatrix}\sigma_n & 0\\ 0 & 1\end{bmatrix}\\
&=\left[\,\varpi^{\alpha *}_{n-1}\phi^{\alpha}_{n-1}\dot{B}^{\beta}_{n-1}\middle|\varpi^{\alpha}_{n}\phi^{\alpha *}_{n}\dot{B}^{\beta}_{n-1}\right]\begin{bmatrix}1 & 0\\ 0 & \overline{\sigma}_n\end{bmatrix}\\
&=\left[\,\varpi^{\alpha *}_{n-1}\phi^{\alpha}_{n-1}\dot{B}^{\beta}_{n-1}\middle|\varpi^{\alpha *}_{n}\phi^{\alpha *}_{n}\dot{B}^{\beta}_{n}\right].
\end{aligned}
$$

Thus

$$
\left[\,\varpi^{\alpha *}_{n-1}\phi^{\alpha}_{n-1}\dot{B}^{\beta}_{n-1}\middle|\varpi^{\alpha *}_{n}\phi^{\alpha *}_{n}\dot{B}^{\beta}_{n}\right]=\left[\,\varpi^{\alpha}_{n-1}\phi^{\alpha *}_{n-1}\dot{B}^{\beta}_{n-1}\middle|\varpi^{\alpha *}_{n}\phi^{\alpha}_{n}\dot{B}^{\beta}_{n}\right]\hat{G}^{\beta}_{n}. \tag{9.4}
$$

We build the general CMV matrix as a product of G-factors. Set $\alpha_0 = 0$ and suppose we consider for $1 \le n < k < m < l < \cdots \le \infty$

$$
\alpha_0, \alpha_1, \ldots, \alpha_{n-1}|\beta_n, \ldots, \beta_{k-1}|\alpha_k, \ldots, \alpha_{m-1}|\beta_m, \ldots, \beta_{\ell-1}|\alpha_\ell, \ldots.
$$

We consider the vector

$$
\begin{aligned}
\Phi(\mathcal{I}-\mathcal{A}^* z) &= \Big[\varpi^{\alpha}_0\phi_0, \ldots, \varpi^{\alpha}_{n-1}\phi_{n-1}\Big|\ \varpi^{\alpha}_{n}\phi_n, \ldots, \varpi^{\alpha}_{k-1}\phi_{k-1}\Big|\ \varpi^{\alpha}_{k}\phi_k, \ldots\\
&\qquad \ldots, \varpi^{\alpha}_{m-1}\phi_{m-1}\Big|\ \varpi^{\alpha}_{m}\phi_m, \ldots, \varpi^{\alpha}_{\ell-1}\phi_{\ell-1}\Big|\ \varpi^{\alpha}_{\ell}\phi_\ell, \ldots]\\
&= \Big[\varpi^{\alpha}_0\phi^{\alpha}_0, \ldots, \varpi^{\alpha}_{n-1}\phi^{\alpha}_{n-1}\dot{B}^{\beta}_{n-1}\Big|\ \varpi^{\alpha}_{n}\phi^{\alpha *}_{n}\dot{B}^{\beta}_{n}, \ldots, \varpi^{\alpha}_{k-1}\phi^{\alpha *}_{k-1}\dot{B}^{\beta}_{k-1}\Big|\ \varpi^{\alpha}_{k}\phi^{\alpha}_{k}\dot{B}^{\beta}_{k}, \ldots\\
&\qquad \ldots, \varpi^{\alpha}_{m-1}\phi^{\alpha}_{m-1}\dot{B}^{\beta}_{m-1}\Big|\ \varpi^{\alpha}_{m}\phi^{\alpha *}_{m}\dot{B}^{\beta}_{m}, \ldots, \varpi^{\alpha}_{\ell-1}\phi^{\alpha *}_{\ell-1}\dot{B}^{\beta}_{\ell-1}\Big|\ \varpi^{\alpha}_{\ell}\phi^{\alpha}_{\ell}\dot{B}^{\beta}_{\ell}, \ldots\Big].
\end{aligned}
$$

Define for $\nu \in \{\alpha, \beta\}$ and with $\eta_{\mathcal{A}} = \sqrt{\mathcal{I}-\mathcal{A}^*\mathcal{A}}$

$$
G^{\nu}_{n} = \begin{bmatrix} I_{n-1} & & \\ & \tilde{G}^{\nu}_{n} & \\ & & I_\infty \end{bmatrix} = \operatorname{diag}(I_{n-1}, \tilde{G}^{\nu}_{n}, I_\infty), \quad \text{and} \quad \hat{G}^{\nu}_{n} = \eta^{-1}_{\mathcal{A}} G^{\nu}_{n} \eta_{\mathcal{A}}. \tag{9.5}
$$

Apply successively the G^{ν}_{n} to the right on the Φ vector. Keep an increasing order for the factors *within* an α-block and a decreasing order for the factors *covering* a β-block. Thus for the example given above the order will be

$$\hat{\mathcal{G}} = \eta_{\mathcal{A}}^{-1}\mathcal{G}\eta_{\mathcal{A}}, \tag{9.6}$$

with

$$\begin{aligned}
\gamma = \quad & \alpha_0, \alpha_1, \ldots, \alpha_{n-1} \,\|\beta_n, \ldots, \beta_{k-1}|\alpha_k\|\alpha_{k+1}, \ldots, \alpha_{m-1}\|\beta_m, \ldots, \beta_{\ell-1}|\alpha_\ell\|\alpha_{\ell+1}, \ldots \\
\hat{\mathcal{G}} = \quad & \underbrace{(\hat{G}_1^\alpha \hat{G}_2^\alpha \cdots \hat{G}_{n-1}^\alpha)}_{\hat{G}_\alpha^1} \underbrace{(\hat{G}_k^\alpha \hat{G}_{k-1}^\beta \cdots \hat{G}_n^\beta)}_{\hat{G}_\beta^1} \underbrace{(\hat{G}_{k+1}^\alpha \cdots \hat{G}_{m-1}^\alpha)}_{\hat{G}_\alpha^2} \underbrace{(\hat{G}_\ell^\alpha \hat{G}_{\ell-1}^\beta \cdots \hat{G}_m^\beta)}_{\hat{G}_\beta^2} (\hat{G}_{\ell+1}^\alpha \cdots) \\
= \eta_{\mathcal{A}}^{-1} & \underbrace{(G_1^\alpha G_2^\alpha \cdots G_{n-1}^\alpha)}_{G_\alpha^1} \underbrace{(G_k^\alpha G_{k-1}^\beta \cdots G_n^\beta)}_{G_\beta^1} \underbrace{(G_{k+1}^\alpha \cdots G_{m-1}^\alpha)}_{G_\alpha^2} \underbrace{(G_\ell^\alpha G_{\ell-1}^\beta \cdots G_m^\beta)}_{G_\beta^2} (G_{\ell+1}^\alpha \cdots)\eta_{\mathcal{A}}
\end{aligned}$$

Using (9.2) when multiplying $\Phi(\mathcal{I} - \mathcal{A}^* z)$ with $\hat{G}_\alpha^1$ gives (note $\phi_0 = \phi_0^*$)

$$\left[\varpi_0^{\alpha*}\phi_0, \ldots, \varpi_{n-2}^{\alpha*}\phi_{n-2}, \varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha*}\dot{B}_{n-1}^{\beta}\right| \varpi_n^\alpha \phi_n, \ldots]$$

While multiplying this result with $\hat{G}_k^\alpha \hat{G}_{k-1}^\beta \cdots \hat{G}_{n+1}^\beta$, we make use of (9.2) and (9.4) to obtain

$$\left[\varpi_0^{\alpha*}\phi_0, \ldots, \varpi_{n-2}^{\alpha*}\phi_{n-2}, \varpi_{n-1}^{\alpha}\phi_{n-1}^{\alpha*}\dot{B}_{n-1}^{\beta}\right| \varpi_n^{\alpha*}\phi_n^\alpha \dot{B}_n, \varpi_{n+1}^{\alpha*}\phi_{n+1}, \ldots, \varpi_{k-1}^{\alpha*}\phi_{k-1}\Big| \varpi_k^\alpha \phi_k^*, \varpi_{k+1}^\alpha \phi_{k+1}, \ldots\Big]$$

and the remaining multiplication $\hat{G}_n^\beta$ fixes the link of the α and β-block:

$$\left[\varpi_0^{\alpha*}\phi_0, \ldots, \varpi_{n-1}^{\alpha*}\phi_{n-1}\right| \varpi_n^{\alpha*}\phi_n, \ldots, \varpi_{k-1}^{\alpha*}\phi_{k-1}\Big| \varpi_k^\alpha \phi_k^*, \varpi_{k+1}^\alpha \phi_{k+1}, \ldots\Big]$$

The next block is again an α-block treated by the product $\hat{G}_\alpha^2$, and one may continue like this to finally get

$$\Phi(\mathcal{I} - \mathcal{A}^* z)\hat{\mathcal{G}} = \Phi(z\mathcal{I} - \mathcal{A}) \tag{9.7}$$

and $\hat{\mathcal{G}} = \eta_{\mathcal{A}}^{-1}\mathcal{G}\eta_{\mathcal{A}}$ with $\mathcal{G}$ the product of unitary matrices, all of which have only one nontrivial 2×2 diagonal block.

Note that for example $G_\beta^2 = G_\ell^\alpha G_{\ell-1}^\beta \cdots G_m^\beta$. Because $\ell > m$ there must be at least two factors in such a β-block: the first one is a $G_\cdot^\alpha$ and the last one is a $G_\cdot^\beta$. However $G_\alpha^2 = G_{k+1}^\alpha \cdots G_{m-1}^\alpha$ thus if $k = m - 1$ then the initial index $k + 1$ is larger than the end index $m - 1$. Thus if an α-block has only one element then there are no factors in this product, which means that this G_α is just the identity.

In the case of $\underline{\boldsymbol{\gamma}} = \underline{\boldsymbol{\alpha}}$ (no β's), then there is of course only one infinite block so $n = \infty$ and there is no $k, m, \ell, \ldots$

$$\mathcal{G} = G_1^\alpha G_2^\alpha G_3^\alpha \cdots = \mathcal{H}$$

is the familiar upper Hessenberg matrix, and in the case of alternating α-β sequence, we have in the case $\alpha_1, \beta_2, \alpha_3, \beta_4, \ldots$

$$\mathcal{G} = G_1^\alpha (G_3^\alpha G_2^\beta) I (G_5^\alpha G_4^\beta) I (G_7^\alpha G_6^\beta) \cdots = (G_1^\alpha G_3^\alpha G_5^\alpha \cdots)(G_2^\beta G_4^\beta G_6^\beta \cdots) = \mathcal{C} = \mathcal{C}_o^\alpha \mathcal{C}_e^\beta$$

because the blocks commute. If we define $\mathcal{S}_e = \operatorname{diag}(1, \tilde{S}_2, \tilde{S}_4, \tilde{S}_6, \ldots)$, $\tilde{S}_{2k} = \operatorname{diag}(\sigma_{2k}, 1)$, then of course $\mathcal{C} = \mathcal{C}_o^\alpha \mathcal{S}_e^* \mathcal{C}_e^\alpha \mathcal{S}_e$.

In the case $\beta_1, \alpha_2, \beta_3, \alpha_4, \ldots$, then

$$\mathcal{G} = (G_2^\alpha G_1^\beta) I (G_4^\alpha G_3^\beta) I (G_6^\alpha G_5^\beta) \cdots = (G_2^\alpha G_4^\alpha G_6^\alpha \cdots)(G_1^\beta G_3^\beta G_5^\beta \cdots) = \mathcal{C} = \mathcal{C}_e^\alpha \mathcal{C}_o^\beta.$$

As in the previous case this is $\mathcal{C} = \mathcal{C}_e^\alpha \mathcal{S}_o^* \mathcal{C}_o^\alpha \mathcal{S}_o$ with $\mathcal{S}_o = \operatorname{diag}(\tilde{S}_1, \tilde{S}_3, \tilde{S}_5, \ldots)$, $\tilde{S}_{2k-1} = \operatorname{diag}(\sigma_{2k-1}, 1)$.

These are the classical CMV matrices as also given in [28], except for the $\mathcal{S}$ factors, which were not in [28]. That is because we have used a particular normalization for the ϕ_n-basis. A slightly different normalization of the ϕ_n-basis will remove the $\mathcal{S}$ factors. Indeed, replacing all ϕ_n by $\varphi_n^\alpha = \dot{\varsigma}_n^\beta \phi_n$, where

$$\dot{\varsigma}_n^\beta = \frac{\dot{B}_n^\beta}{|\dot{B}_n^\beta|} = \prod_{j \in \mathbb{b}_n} \sigma_j,$$

will do.

To see this note that the relation (9.3) between $\tilde{G}_n^\beta$ and $\tilde{G}_n^\alpha$ can also be written as (multiply with $\sigma_n \overline{\sigma}_n = 1$)

$$\tilde{G}_n^\beta = \begin{bmatrix} 1 & 0 \\ 0 & \sigma_n \end{bmatrix} \tilde{G}_n^\alpha \begin{bmatrix} 1 & 0 \\ 0 & \overline{\sigma}_n \end{bmatrix}. \tag{9.8}$$

The first and the last of these factors can then be moved to the φ_n so that the β-relation (9.4) now becomes

$$\left[\varpi_{n-1}^\alpha \dot{\varsigma}_{n-1}^\beta \phi_{n-1}^\alpha \dot{B}_{n-1}^\beta \middle| \varpi_n^{\alpha*} \dot{\varsigma}_n^\beta \phi_n^{\alpha*} \dot{B}_n^\beta \right] = \left[\varpi_{n-1}^\alpha \dot{\varsigma}_{n-1}^\beta \phi_{n-1}^{\alpha*} \dot{B}_{n-1}^\beta \middle| \varpi_n^{\alpha*} \dot{\varsigma}_n^\beta \phi_n^\alpha \dot{B}_n^\beta \right] \hat{G}_n^\alpha. \tag{9.9}$$

Note that the multiplication on the right is with $\hat{G}_n^\alpha$ and not with $\hat{G}_n^\beta$.

For the α-case, i.e., $\gamma_n = \alpha_n$ nothing essentially changes since then $\dot{\varsigma}_n^\beta = \dot{\varsigma}_{n-1}^\beta$.

It should be clear that following the same arguments used above, the relation (9.7) then becomes

$$\Phi^\alpha(\mathcal{I}-\mathcal{A}^*z)\hat{\mathcal{G}}^\alpha=\Phi^\alpha(z\mathcal{I}-\mathcal{A})\quad\text{with}\quad\Phi^\alpha=[\varphi_0^\alpha,\varphi_1^\alpha,\varphi_2^\alpha,\ldots]$$

and $\hat{\mathcal{G}}^\alpha$ is exactly like $\hat{\mathcal{G}}$ in (9.6), except that all G_j^β should be replaced by a G_j^α.

From (9.7) we derive that

$$z\mathcal{I}=(\hat{\mathcal{G}}+\mathcal{A})(\mathcal{I}+\mathcal{A}^*\hat{\mathcal{G}})^{-1}=\eta_{\mathcal{A}}^{-1}(\mathcal{G}+\mathcal{A})(\mathcal{I}+\mathcal{A}^*\mathcal{G})^{-1}\eta_{\mathcal{A}}$$

which is the matrix representation of the shift operator with respect to the basis $(\phi_0,\phi_1,\phi_2,\ldots)$. With respect to the basis Φ^α, the expression is the same except that $\mathcal{G}$ should be replaced by $\mathcal{G}^\alpha$.

Of course the spectrum of the operator will not be affected by the renormalization factor $\dot{\sigma}_n^\beta$, i.e., the spectrum remains the same whether or not the $\mathcal{S}$ factors are present.

We summarize our previous results.

Theorem 9.2 *In the general case of a sequence $\underline{\boldsymbol{\gamma}}$, then with the notation introduced above, the shift operator with respect to the orthonormal basis $\{\phi_n\}$ has the form*

$$z\mathcal{I}=(\hat{\mathcal{G}}+\mathcal{A})(\mathcal{I}+\mathcal{A}^*\hat{\mathcal{G}})^{-1}=\eta_{\mathcal{A}}^{-1}(\mathcal{G}+\mathcal{A})(\mathcal{I}+\mathcal{A}^*\mathcal{G})^{-1}\eta_{\mathcal{A}}$$

where $\mathcal{A}=\operatorname{diag}(\alpha_0,\alpha_1,\ldots)$, $\eta_{\mathcal{A}}=\sqrt{\mathcal{I}-\mathcal{A}^*\mathcal{A}}$ and $\mathcal{G}$ is a product of unitary factors G_k defined in (9.5) where the order of multiplication is from left to right in a block of successive α-values and from right to left in a block of successive β-values as explained in detail above.
For the sequence $\underline{\boldsymbol{\alpha}}$ we get the classical Hessenberg matrix

$$\mathcal{G}=G_1^\alpha G_2^\alpha G_3^\alpha\cdots=\mathcal{H}.$$

The classical CMV matrices are obtained for $\alpha_1,\beta_2,\alpha_3,\beta_4,\ldots$ as

$$\mathcal{G}=G_1^\alpha(G_3^\alpha G_2^\beta)(G_5^\alpha G_4^\beta)(G_7^\alpha G_6^\beta)\cdots=(G_1^\alpha G_3^\alpha G_5^\alpha\cdots)(G_2^\beta G_4^\beta G_6^\beta\cdots)=\mathcal{C}=\mathcal{C}_o^\alpha\mathcal{C}_e^\beta$$

and in the case $\beta_1,\alpha_2,\beta_3,\alpha_4,\ldots$, then

$$\mathcal{G}=(G_2^\alpha G_1^\beta)(G_4^\alpha G_3^\beta)(G_6^\alpha G_5^\beta)\cdots=(G_2^\alpha G_4^\alpha G_6^\alpha\cdots)(G_1^\beta G_3^\beta G_5^\beta\cdots)=\mathcal{C}=\mathcal{C}_e^\alpha\mathcal{C}_o^\beta.$$

It we use the slightly different orthonormalized basis $\{\varphi_n^\alpha\}$, then the previous relations still hold true, except that all G_j^β can be replaced by G_j^α.

Note that $\lambda_n^\alpha = 0$ does not give any problem for these formulas of the G-factors. If $\underline{\gamma} = \underline{\boldsymbol{\alpha}}$ then all ϕ_n^α are regular, but for a general sequence $\underline{\gamma}$, it is possible that ϕ_n is not regular. Recall that then $\lambda_n = \infty$ but by Theorem 5.3 this means that $\lambda_n^\alpha = 0$ and thus there is no problem in using the G-matrices introduced above, even if the ORF sequence is not regular.

Another thing to note here is that, as we remarked before, the factors G_k^ν are unitary, although the factors $\hat{G}_k^\nu = \eta_{\mathcal{A}}^{-1} G_k^\nu \eta_{\mathcal{A}}$ are in general not unitary. However $\tilde{\zeta}_{\mathcal{A}}(\mathcal{G}) = \eta_{\mathcal{A}}^{-1}(\mathcal{G}+\mathcal{A})(\mathcal{I}+\mathcal{A}^*\mathcal{G})^{-1}\eta_{\mathcal{A}}$ is unitary when $\mathcal{G}$ is unitary as can be directly verified in the finite dimensional case. For example $[\tilde{\zeta}_{\mathcal{A}}(\mathcal{G})][\tilde{\zeta}_{\mathcal{A}}(\mathcal{G})]^* = \mathcal{I}$ if and only if

$$(\mathcal{I} + \mathcal{A}\mathcal{G}^*)\eta_{\mathcal{A}}^{-2}(\mathcal{I} + \mathcal{G}\mathcal{A}) = (\mathcal{G}^* + \mathcal{A}^*)\eta_{\mathcal{A}}^{-2}(\mathcal{G} + \mathcal{A})$$

which is verified after working out and using that $\mathcal{A}$ and $\eta_{\mathcal{A}}$ commute. A similar calculation verifies that $[\tilde{\zeta}_{\mathcal{A}}(\mathcal{G})]^*[\tilde{\zeta}_{\mathcal{A}}(\mathcal{G})] = \mathcal{I}$. The same holds for the infinite case i.e, in the whole of L_μ^2 if $\mathcal{L}$ is dense (see also [28]).

When using the recursion for the ϕ_n with respect to a general sequence $\underline{\gamma}$, then one will not directly compute the λ_n^α, which requires the use of the sequence $\underline{\boldsymbol{\alpha}}$ but whenever λ_n is computed, then again by this Theorem 5.3, we can derive from these the λ_n^α. Of course if we know the sequence $\underline{\gamma}$, then we also know the sequence $\underline{\boldsymbol{\alpha}}$. So we can always generate the matrix G_n as fully expressed in terms of α-related quantities. Equivalently, one may express the previous matrix G_n^α in terms of the λ_n. There is only one problem: the computation of η_{n1}^α for which we did not give and explicit expression so far. So we prove:

Lemma 9.3 *The phase θ_n^α of the unitary factor $\eta_{n1}^\alpha = e^{i\theta_n^\alpha}$ is given by*

$$\theta_n^\alpha = \arg\Big(\overline{\sigma}_{n-1}\sigma_n\overline{\varpi_n^\alpha(\alpha_{n-1})\phi_n^{\alpha*}(\alpha_{n-1})}\Big) \quad \text{or} \quad \eta_{n1}^\alpha = \overline{\sigma}_{n-1}\sigma_n\mathfrak{u}\Big(\varpi_n^\alpha(\alpha_{n-1})\phi_n^{\alpha*}(\alpha_{n-1})\Big).$$

(Recall $\mathfrak{u}(z) = \overline{z}/|z|$.)

PROOF. Take the relation (9.1) and evaluate for $z = \alpha_{n-1}$, then the left-hand side vanishes because $\varpi_{n-1}^{\alpha*}(\alpha_{n-1}) = 0$. Hence

$$\lambda_n^\alpha\eta_{n1}^\alpha\frac{\varpi_{n-1}^\alpha(\alpha_{n-1})\phi_{n-1}^{\alpha*}(\alpha_{n-1})}{\sqrt{1-|\alpha_{n-1}|^2}} = \sqrt{1-|\lambda_n^\alpha|^2}\frac{\varpi_n^\alpha(\alpha_{n-1})\phi_n^\alpha(\alpha_{n-1})}{\sqrt{1-|\alpha_n|^2}}.$$

or

$$\eta_{n1}^\alpha = \frac{\sqrt{1-|\lambda_n^\alpha|^2}}{\sqrt{1-|\alpha_{n-1}|^2}\sqrt{1-|\alpha_n|^2}}\frac{\varpi_n^\alpha(\alpha_{n-1})\phi_n^\alpha(\alpha_{n-1})}{\lambda_n^\alpha\phi_n^{\alpha*}(\alpha_{n-1})}.$$

Use the definition of $\lambda_n^\alpha = \eta_n^\alpha \frac{\phi_n^\alpha(\alpha_{n-1})}{\phi_n^{\alpha*}(\alpha_{n-1})}$ with $\eta_n^\alpha = \sigma_{n-1}\overline{\sigma}_n \frac{\varpi_n^\alpha(\alpha_{n-1})}{\varpi_{n-1}^\alpha(\alpha_n)}$ and knowing that $\phi_{n-1}^{\alpha*}(\alpha_{n-1}) > 0$ we obtain after simplification and leaving out all the factors with phase zero

$$\theta_n^\alpha = \arg\left(\overline{\sigma}_{n-1}\sigma_n\overline{\varpi_n^\alpha(\alpha_{n-1})\phi_n^{\alpha*}(\alpha_{n-1})}\right)$$

as claimed. $\square$

Note that this expression for η_{n1}^α is well defined because $\phi_n^{\alpha*}(\alpha_{n-1}) \neq 0$.

We are now ready to show how to use the quantities for a general $\underline{\gamma}$-sequence in generating the G_n^α matrix. For example if $n-1 \in \mathbb{b}_n$ and $n \in \mathbb{a}_n$, then $\gamma_n = \alpha_n$, $\gamma_{n-1} = \beta_{n-1}$, $\lambda_n^\alpha = 1/\overline{\lambda}_n$, $\phi_n^\alpha = \phi_n/\dot{B}_n^\beta$. This allows to express η_{n1}^α and all the other elements of G_n in terms of the $\underline{\gamma}$-related elements. If we assume that ϕ_n is regular, then

$$\begin{aligned}\tilde{G}_n^\alpha &= \overline{\sigma}_{n-1}\overline{\eta}_{n1}^\alpha \begin{bmatrix} -\eta_{n1}^\alpha/\overline{\lambda}_n & \sqrt{1-1/|\lambda_n|^2} \\ \sqrt{1-1/|\lambda_n|^2} & \overline{\eta}_{n1}^\alpha/\lambda_n \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sigma_n \end{bmatrix} \\ &= \frac{\overline{\sigma}_{n-1}\overline{\eta}_{n1}^\alpha}{|\lambda_n|} \begin{bmatrix} u_n & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & \sqrt{|\lambda_n|^2-1} \\ \sqrt{|\lambda_n|^2-1} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \overline{u}_n\sigma_n \end{bmatrix}\end{aligned}$$

where $u_n = (\lambda_n/|\lambda_n|)\eta_{n1}^\alpha \in \mathbb{T}$.

Expressing η_{n1}^α explicitly in terms of $\underline{\gamma}$-elements is a bit messy. Suppose that $1/\overline{\alpha}_{n-1} = \beta_{n-1} \in \mathbb{b}_n$, then ϕ_n^* has a pole in α_{n-1} although this is compensated by the same pole of $\dot{B}_n^\beta$ in $\phi_n^{\alpha*} = \phi_n^*/\dot{B}_n^\beta$. Indeed as shown in (3.1)

$$\begin{aligned}\phi_n^{\alpha*}(z) &= \frac{\phi_n^*(z)}{\dot{B}_n^\beta(z)} = \dot{B}_n^\alpha(z)\phi_{n*}(z) = \frac{\dot{\varsigma}_n^\alpha\, p_n^*(z)}{\dot{\pi}_n^\alpha(z)\prod_{j\in\mathbb{b}_n}(z-\beta_j)}, \quad \dot{\varsigma}_n^\alpha = \prod_{j\in\mathbb{a}_n}\sigma_j \\ &= \frac{p_n^*(z)\dot{\varsigma}_n^\alpha\prod_{j\in\mathbb{b}_n}(-\overline{\alpha}_j)}{\pi_n^\alpha(z)} = \frac{\varsigma_n p_n^*(z)\prod_{j\in\mathbb{b}_n}|\alpha_j|}{\pi_n^\alpha(z)} = \frac{\varsigma_n p_n^*(z)}{\pi_n^\alpha(z)\prod_{j\in\mathbb{b}_n}|\beta_j|}\end{aligned}$$

which can always be evaluated for $z = \alpha_{n-1}$ without a problem. Thus evaluate the numerator p_n^* of ϕ_n^* at $z = \alpha_{n-1}$, multiply by ς_n and divide by the product of $\pi_{n-1}^\alpha(\alpha_{n-1})$ and $\prod_{j\in\mathbb{b}_n}|\beta_j|$ and this will give you the value of $\varpi_n^\alpha(\alpha_{n-1})\phi_n^{\alpha*}(\alpha_{n-1})$, which will define the phase for η_{n1}^α as given in Lemma 9.3.

If n is not a regular index, then $\lambda_n = \infty$ and in that case the matrix $\tilde{G}_n^\alpha$ has the form

$$\tilde{G}_n^\alpha = \overline{\sigma}_{n-1}\overline{\eta}_{n1}^\alpha \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sigma_n \end{bmatrix} = \overline{\sigma}_{n-1}\overline{\eta}_{n1}^\alpha \begin{bmatrix} 0 & \sigma_n \\ 1 & 0 \end{bmatrix}.$$

10 Spectral analysis

Recall the operator Möbius transform of the previous section. If $\|\mathcal{A}\| < 1$ then $\tilde{\zeta}_{\mathcal{A}}(\mathcal{T})$ will be isometric/unitary whenever $\mathcal{T}$ is isometric/unitary [28]. Thus if the Hessenberg or the CMV matrix or general $\mathcal{G}$ matrix is unitary, then also its Möbius transform will be unitary.

Now [28, Thm. 4.2] says that if $\underline{\boldsymbol{\alpha}}$ is compactly included in $\mathbb{D}$ and μ satisfies the Szegő condition $\mu' \notin L^1$, then $\mathcal{V} = \tilde{\zeta}_{\mathcal{A}}(\mathcal{H})$ will have $\operatorname{supp}\mu = \sigma(\mathcal{V})$ and a mass point λ of μ is a simple eigenvalue of $\mathcal{V}$ with eigenvector $\Phi^\alpha(\lambda) = (\phi_0^\alpha(\lambda), \phi_1^\alpha(\lambda), \phi_2^\alpha(\lambda), \ldots)$ and $\mu(\{\lambda\}) = 1/\|\Phi^\alpha(\lambda)\|^2$. Recall that whatever $\underline{\boldsymbol{\gamma}}$, we have chosen the ORF ϕ_n to be orthonormal, an hence $\phi_0 = 1$.

Similarly [28, Thm. 5.1, 5.4] says that if $\underline{\boldsymbol{\alpha}}$ is compactly included in $\mathbb{D}$, and $\underline{\boldsymbol{\alpha}} = \underline{\boldsymbol{\varepsilon}}$ is the alternating sequence mentioned above, then $\mathcal{U} = \tilde{\zeta}_{\mathcal{A}}(\mathcal{C})$ will be the matrix representation of the shift $\mathcal{T}_\mu$ with respect to the basis $(\phi_n^\varepsilon)_{n\geq 0}$ and $\operatorname{supp}\mu = \sigma(\mathcal{U})$ and a mass point λ of μ is a simple eigenvalue of $\mathcal{U}$ with eigenvector $\Phi^\varepsilon(\lambda) = (\phi_0^\varepsilon(\lambda), \phi_1^\varepsilon(\lambda), \phi_2^\varepsilon(\lambda), \ldots)$ and $\mu(\{\lambda\}) = 1/\|\Phi^\alpha(\lambda)\|^2 = 1/\|\Phi^\varepsilon(\lambda)\|^2$.

Thus if $\mathcal{T}_\mu$ is the unitary shift operator in L^2_μ, i.e., $\mathcal{T}_\mu f(z) = zf(z)$ for all $f \in L^2_\mu$, then the previous relations say that in the case $\underline{\boldsymbol{\gamma}} = \underline{\boldsymbol{\alpha}}$ or $\underline{\boldsymbol{\gamma}} = \underline{\boldsymbol{\varepsilon}}$ then $\mathcal{T}_\mu \restriction \mathcal{L}$ with $\mathcal{L} = \operatorname{cl}_{L^2_\mu} \operatorname{span}\{\phi_0^\nu, \phi_1^\nu, \ldots\}$, has an isometric matrix representation in the ϕ^ν-basis given by $\tilde{\zeta}_{\mathcal{A}}(\mathcal{G}^\nu)$ for $\nu \in \{\alpha, \varepsilon\}$. If $\underline{\boldsymbol{\alpha}}$ is compactly included in $\mathbb{D}$, then $(\phi_k^\alpha)_{k\geq 0}$ is complete in L^2_μ, and $\tilde{\zeta}_{\mathcal{A}}(\mathcal{G}^\alpha) = \tilde{\zeta}_{\mathcal{A}}(\mathcal{H})$ is unitary and represents the full operator.

Under the same conditions also $(\phi_k^\varepsilon)_{k\geq 0}$ is a basis for L^2_μ and $\tilde{\zeta}_{\mathcal{A}}(\mathcal{G}^\varepsilon)$ is unitary and represents the full shift operator. In fact, [28, Thm. 5.3] says for the sequence $\underline{\boldsymbol{\varepsilon}}$ that $\sum_{k=1}^\infty (1 - |\alpha_{2k}|^2) = \sum_{k=1}^\infty (1 - |\alpha_k|^{2k-1}) = \infty$ is sufficient for both $(\phi_{2k-1}^\varepsilon = \phi_{2k-1}^\alpha)_{k\geq 1}$ and $(\phi_{2k}^\varepsilon = \phi_{2k*}^\beta)_{k\geq 0}$ to form a basis for L^2_μ.

All this suggests that in the general case $\tilde{\zeta}_{\mathcal{A}}(\mathcal{G})$ will be a representation with respect to the general set $(\phi_k)_{k\geq 0}$ of $\mathcal{T}_\mu \restriction \mathcal{L}$ with $\mathcal{L} = \operatorname{cl}_{L^2_\mu} \operatorname{span}\{\phi_0, \phi_1, \ldots\}$

It is an easy adaptation of [28, Proposition 5.3] to find that $\mathcal{L} = L^2_\mu$ if $\sum_{k\in \mathrm{a}_\infty} (1 - |\alpha_k|^2) = \infty$

and $\sum_{k\in\mathbb{b}_\infty}(1-|\alpha_k|^2)=\infty$. In that case $\tilde{\zeta}_{\mathcal{A}}(\mathcal{G})$ is the representation of the full operator.

Of course the spectrum of the operator will not be affected by the renormalization factor $\dot{\varsigma}_n^\beta$ that we used in the previous section, i.e., the spectrum remains the same whether or not the $\mathcal{S}$ factors are present. This means that, on condition that $\mathcal{L}$ is dense in L^2_μ, we work using only the G_n^α factors in the factorization of the operator $\mathcal{G}$, and the spectrum will not be affected by the order in which these are multiplied.

This is a well known fact in the case of a finite unitary Hessenberg matrix. It can be decomposed as a finite product of finite matrices of the type $\mathcal{H}=\hat{G}_1\hat{G}_2\cdots\hat{G}_n$. Multiplying these factors in any order will not change the spectrum of the matrix. See e.g. [1]. The previous discussion illustrates that the same observation holds for the infinite case and this transfers to the rational case after the matrix Möbius transform as well. The shape of the general $\mathcal{G}$ matrix is also implicit in the computations for the rational Krylov method in [23].

In fact, this property about the spectrum holds for general AMPD matrices. AMPD matrices are of the form $AM+D$ with A and D arbitrary diagonal matrices and M a product of G_k-type factors, i.e., these are the identity with only an arbitrary 2×2 diagonal block on position k. The spectrum of an AMPD matrix does not depend on the order in which the G_k factors are multiplied. The same property holds for rational forms, i.e., the spectrum of $(BM+E)(AM+D)^{-1}$ with $AM+D$ and $BM+E$ both AMPD matrices, will be independent of the order in which the product defining M is organized. This will be proved in a separate forthcoming paper. The present application of this property is a special case since it requires that all the matrices involved are unitary. Then the additional property holds that although the eigenvectors of the matrix will depend on M, i.e., the order of multiplication, the absolute values of the components of the eigenvectors turn out to be the same since for a given eigenvalue $\lambda\in\mathbb{T}$ they are given by $|\phi_k(\lambda)|=|\phi_k^\alpha(\lambda)|=|\phi_k^\beta(\lambda)|$.

11 Computation of the quadrature

To construct the nodes of the n-point quadrature formulas via an eigenvalue decomposition of a unitary truncation of matrix $\tilde{\zeta}_{\mathcal{A}}(\mathcal{C})$ in the general case is not a good option for practical use, because truncating to $\mathcal{L}_n$ is much more complicated because the Hessenberg blocks are on both sides of the diagonal. However because the order in which the poles are introduced does not play a role in the eventual quadrature formula. Hence the most economic solution is to take the sparsest representation, which is the CMV matrix corresponding to the alternating case. In that case the nodes and weights of the quadrature formula can be computed as described in [3, Thm. 7.3, Cor. 8.3, Prop. 8.4]. Note that we never build a Hessenberg block. What we have to do is, no matter what sequence $\underline{\boldsymbol{\gamma}}$ is, we always take the underlying $\underline{\boldsymbol{\alpha}}$ sequence and we alternate choosing an α_k followed by a $\beta_k = 1/\overline{\alpha}_{k+1}$, so that we have the CMV structure. The space in which the n-point quadrature formula will be exact is always the same: $\mathcal{R}_{n-1} = \mathcal{L}_{n-1} \cdot \mathcal{L}_{(n-1)*}$.

Rational interpolatory and Szegő quadrature formulas with arbitrary order of the poles including error estimates and numerical experiments were also considered in [10].

12 Alternative formulations

12.1 Using the general recursion and dealing with infinity

Suppose $\underline{\boldsymbol{\gamma}} = \underline{\boldsymbol{\alpha}}$, then our derivation given in sections 8 and 9 shows that

$$\Phi(\mathcal{I} - \mathcal{A}^* z)\hat{\mathcal{H}}_\alpha = \Phi(z\mathcal{I} - \mathcal{A})$$

with $\hat{\mathcal{H}}_\alpha$ an upper Hessenberg matrix. If we go through the derivation, then the same arguments used for the sequence $\underline{\boldsymbol{\alpha}}$ also applies when we do exactly the same steps using the recursion of Theorem 5.1 for the general sequence $\underline{\boldsymbol{\gamma}}$, where we assume for the moment that $\gamma_i = \infty$ does not appear. Then we shall again arrive at the above relation, except that all α's are replaced by γ's. Thus (assume for simplicity that the whole sequence is regular) and recall that all $|\gamma_k| \neq 1$ then

$$\Phi(\mathcal{I} - \mathit{\Gamma}^* z)\hat{\mathcal{H}}_\gamma = \Phi(z\mathcal{I} - \mathit{\Gamma}) \quad \text{and} \quad z\mathcal{I} = (\hat{\mathcal{H}}_\gamma + \mathit{\Gamma})(\mathcal{I} + \mathit{\Gamma}^*\hat{\mathcal{H}}_\gamma)^{-1} \tag{12.1}$$

where $\mathit{\Gamma} = \operatorname{diag}(\gamma_0, \gamma_1, \ldots)$ and $\eta_\gamma = (\mathcal{I} - \mathit{\Gamma}^*\mathit{\Gamma})^{1/2}$ in $\hat{\mathcal{H}}_\gamma = \eta_\gamma^{-1}\mathcal{H}_\gamma\eta_\gamma$ which is again upper Hessenberg. An important note is in order here. The previous notation is purely formal. The practical expressions will involve quatities $1 - |\lambda_k|^2$ and $1 - |\gamma_k|^2$ that can be negative so that for example the definition of η_γ is a bit problematic. However, remember the relation (5.1), which was the consequence of the fact that $1 - |\lambda_n|^2$ can only be negative if $(1 - |\gamma_n|^2)(1 - |\gamma_{n-1}|^2) < 0$. Therefore bringing the factor η_γ and η_γ^{-1} inside the G-factors will give square roots of positive elements. Just note that

$$
\begin{bmatrix} 1/\sqrt{1-|\gamma_{n-1}|^2} & \\ & 1/\sqrt{1-|\gamma_n|^2} \end{bmatrix}
\begin{bmatrix} -\lambda_n\eta_{n1} & \sqrt{1-|\lambda_n|^2} \\ \sqrt{1-|\lambda_n|^2} & \overline{\lambda}_n\overline{\eta}_{n1} \end{bmatrix}
\begin{bmatrix} \sqrt{1-|\gamma_{n-1}|^2} & \\ & \sqrt{1-|\gamma_n|^2} \end{bmatrix}
$$

$$
= \begin{bmatrix} -\lambda_n\eta_{n1} & \sqrt{\frac{(1-|\lambda_n|^2)(1-|\gamma_n|^2)}{(1-|\gamma_{n-1}|^2)}} \\ \sqrt{\frac{(1-|\lambda_n|^2)(1-|\gamma_{n-1}|^2)}{(1-|\gamma_n|^2)}} & \overline{\lambda}_n\overline{\eta}_{n1} \end{bmatrix}
$$

and all square roots are taken from positive numbers. To keep the link with what was done for the $\underline{\boldsymbol{\alpha}}$ sequence, we used the previous notation and we shall continue doing that in the sequel. In any case the representation $\hat{\mathcal{H}}_\gamma$ is a Hessenberg matrix that can can be computed in a proper way.

The pole at the origin is excluded because it would require that the diagonal elements $\varpi_k(z) = 1 - \overline{\gamma}_k z$ to be replaced by $-z$ if $\gamma_k = \infty$ and similarly $\varpi_k^*(z) = z - \gamma_k$ by -1. To avoid this, we can use the strategy that was used before when switching between the formalism for the $\underline{\boldsymbol{\alpha}}$ and for the $\underline{\boldsymbol{\beta}}$ sequence, except that we now use this strategy to switch between a finite and an infinite entry in the $\underline{\boldsymbol{\gamma}}$ sequence.

Without going through all the details, the duality between $\underline{\boldsymbol{\alpha}}$ and $\underline{\boldsymbol{\beta}}$ can be transferred to a duality between $\underline{\boldsymbol{\gamma}} = \{\gamma_1, \gamma_2, \ldots\} \subset \hat{\mathbb{C}} \setminus \mathbb{T}$ and the reciprocal sequence $\underline{\check{\boldsymbol{\gamma}}} = \{\check{\gamma}_1, \check{\gamma}_2, \ldots\}$ with $\check{\gamma}_k = 1/\overline{\gamma}_k$, $k = 1, 2, \ldots$. The ORF for the sequence $\underline{\boldsymbol{\gamma}}$ we denote as ϕ_n and the ORF for the sequence $\underline{\check{\boldsymbol{\gamma}}}$ we denote as $\check{\phi}_n$. If $\underline{\boldsymbol{\nu}} = \{\nu_1, \nu_2, \ldots\}$ is a sequence that picks $\nu_k = \gamma_k$ or $\nu_k = \check{\gamma}_k$, then, in analogy with Theorem 3.8, with proper normalization, the ORF for this $\underline{\boldsymbol{\nu}}$ sequence, which are denoted as ϕ_k^ν will be given by

$$
\phi_n^\nu = \phi_n \check{B}_n \quad \text{if} \quad \nu_n = \gamma_n \quad \text{and} \quad \phi_n^\nu = (\phi_n)^* \check{B}_n \quad \text{if} \quad \nu_n = \check{\gamma}_n,
$$

where

$$
\check{B}_n = \prod_{\substack{k=1 \\ \nu_k = \check{\gamma}_k}}^{n} \zeta_k^\nu .
$$

First note that as long as we are dealing with finite γ_k's we can use the recurrence relation and write the analog of the relation given in Theorem 9.1, just replacing the α's by γ's (which by our convention means to remove the superscripts) and write

$$
\left[\varpi_{n-1}^*\phi_{n-1} | \varpi_n \phi_n^*\right] = \left[\varpi_{n-1}\phi_{n-1}^* | \varpi_n \phi_n\right] \hat{G}_n^\gamma \tag{12.2}
$$

with

$$\hat{G}_n^\gamma = (N_n^\gamma)^{-1}\tilde{G}_n^\gamma(N_n^\gamma), \quad N_n^\gamma = \begin{bmatrix} (1-|\gamma_{n-1}|^2)^{-1/2} & 0 \\ 0 & (1-|\gamma_n|^2)^{-1/2} \end{bmatrix},$$

$$\tilde{G}_n^\gamma = \overline{\sigma}_{n-1}\overline{\eta}_{n1} \begin{bmatrix} -\lambda_n\eta_{n1} & \sqrt{1-|\lambda_n|^2} \\ \sqrt{1-|\lambda_n|^2} & \overline{\lambda_n}\overline{\eta}_{n1} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sigma_n \end{bmatrix}$$

Now, if some $\gamma_k = \infty$ is involved, we shall switch to the $\check{\underline{\gamma}}$ sequence, which means that we avoid $\gamma_k = \infty$ and replace it by $\check{\gamma}_k = 0$. The factor $\check{B}_n^\gamma$ will thus only pick up a $\zeta_k^\nu = 1/z$ each time a $\gamma_k = \infty$, hence $\nu_k = 0$.

$$\check{B}_n(z) = \prod_{\substack{j=1 \\ \gamma_j=\infty}}^{n} z = \prod_{j\in\mathbb{i}_n} z = z^{|\mathbb{i}|}$$

where $\mathbb{i}_n = \{j : \gamma_j = \infty, 1 \le j \le n\}$. Since in the previous relations, as long as $\gamma_n \neq \infty$, we have $\check{B}_{n-1} = \check{B}_n$ so that it is no problem to write

$$\left[\varpi_{n-1}\check{B}_{n-1}\phi_{n-1}|\varpi_n\check{B}_n\phi_n^*\right] = \left[\varpi_{n-1}\check{B}_{n-1}\phi_{n-1}^*|\varpi_n\check{B}_n\phi_n\right]\hat{G}_n^\nu$$

with $\hat{G}_n^\nu = \hat{G}_n^\gamma$. If however $\gamma_n = \infty$, then $\varpi_n = -z$ and this z-factor we absorb in $\check{B}_n$. This is to say that $\varpi_n\check{B}_{n-1} = \varpi_n^*\check{B}_n$. Hence

$$\check{B}_{n-1}\left[\varpi_{n-1}\phi_{n-1}^*|\varpi_n\phi_n\right]\hat{G}_n^\nu = \check{B}_{n-1}\left[\varpi_{n-1}\phi_{n-1}|\varpi_n\phi_n^*\right] = \left[\varpi_{n-1}\check{B}_{n-1}\phi_{n-1}|\varpi_n^*\check{B}_n\phi_n^*\right] \tag{12.3}$$

but now $\hat{G}_n^\nu$ is like $\hat{G}_n^\gamma$ but with $\gamma_n = \infty$ replaced by 0. We can proceed exactly as in the case of the α-β duality, but now apply it to the finite-infinite duality to find that

$$\hat{\mathcal{G}} = \eta_{\mathcal{P}}^{-1}\mathcal{G}\eta_{\mathcal{P}}, \tag{12.4}$$

with (all the γ_i's denoted explicitly as such are finite values and denoted as ∞ when not finite)

$$\begin{aligned}
\gamma = \quad & \gamma_0, \gamma_1, \ldots, \gamma_{n-1} \,\|\, \infty, \ldots, \infty | \, \gamma_k \,\|\, \gamma_{k+1}, \ldots, \gamma_{m-1} \| \, \infty, \ldots, \infty | \, \gamma_\ell \| \gamma_{\ell+1}, \ldots \\
\hat{\mathcal{G}} = \quad & \underbrace{(\hat{G}_1^\gamma\hat{G}_2^\gamma\cdots\hat{G}_{n-1}^\gamma)}_{\hat{G}_\gamma^1}\underbrace{(\hat{G}_k^\gamma\hat{G}_{k-1}^\nu\cdots\hat{G}_n^\nu)}_{\hat{G}_\infty^1}\underbrace{(\hat{G}_{k+1}^\gamma\cdots\hat{G}_{m-1}^\gamma)}_{\hat{G}_\gamma^2}\underbrace{(\hat{G}_\ell^\gamma\hat{G}_{\ell-1}^\nu\cdots\hat{G}_m^\nu)}_{\hat{G}_\infty^2}(\hat{G}_{\ell+1}^\gamma\cdots) \\
= \eta_{\mathcal{P}}^{-1} & \underbrace{(G_1^\gamma G_2^\gamma\cdots G_{n-1}^\gamma)}_{G_\gamma^1}\underbrace{(G_k^\gamma G_{k-1}^\nu\cdots G_n^\nu)}_{G_\infty^1}\underbrace{(G_{k+1}^\gamma\cdots G_{m-1}^\gamma)}_{G_\gamma^2}\underbrace{(G_\ell^\gamma G_{\ell-1}^\nu\cdots G_m^\nu)}_{G_\infty^2}(G_{\ell+1}^\gamma\cdots)\eta_{\mathcal{P}}.
\end{aligned}$$

Thus in the γ-blocks we multiply the successive factors G_i^γ to the right while in an ∞-block we multiply in reversed order like we did in the α-β case.

Use (12.2) and multiply $\Phi^\nu(\mathcal{I}-\mathcal{N}^*z)$ from the right with with $\hat{G}^1_\gamma$ and $\mathcal{N}=\mathrm{diag}(\nu_0,\nu_1,\nu_2,\ldots)$ where $\nu_k=0$ if $\gamma_k=\infty$ and $\nu_k=\gamma_k$ otherwise. This gives (note $\phi^\nu_0=\phi^{\nu*}_0$ and $\check{B}_{n-1}=1$)

$$\left[\varpi_0^*\phi_0,\ldots,\varpi_{n-2}^*\phi_{n-2},\varpi_{n-1}\phi_{n-1}^*\check{B}_{n-1}\middle|\ \varpi_n\phi_n^*\check{B}_n,\ldots\right].$$

While multiplying this result with $\hat{G}^\gamma_k\hat{G}^\gamma_{k-1}\cdots\hat{G}^\gamma_{n+1}$, we make use of (12.2) and (12.3) to obtain

$$\left[\varpi_0^*\phi_0,\ldots,\varpi_{n-2}^*\phi_{n-2},\varpi_{n-1}\phi_{n-1}^*\check{B}_{n-1}\middle|\ \varpi_n^*\phi_n\check{B}_n,\varpi_{n+1}^*\phi_{n+1},\ldots,\varpi_{k-1}^*\phi_{k-1}\middle|\ \varpi_k\phi_k^*,\varpi_{k+1}\phi_{k+1},\ldots\right]$$

and the remaining multiplication $\hat{G}^\gamma_n$ fixes the link of the γ and ∞-block, so that after reintroducing the ϕ^ν_i notation:

$$\left[\varpi_0^*\phi_0^\nu,\ldots,\varpi_{n-1}^*\phi_{n-1}^\nu\middle|\ \varpi_n^*\phi_n^\nu,\ldots,\varpi_{k-1}^*\phi_{k-1}^\nu\middle|\ \varpi_k\phi_k^{\nu*},\varpi_{k+1}\phi_{k+1}^\nu,\ldots\right].$$

The next block is again an γ-block treated by the product $\hat{G}^2_\gamma$, and one may continue like this to finally get

$$\Phi^\nu(\mathcal{I}-\mathcal{N}^*z)\hat{\mathcal{G}}=\Phi^\nu(z\mathcal{I}-\mathcal{N}). \tag{12.5}$$

12.2 Alternative not using the matrix Möbius transform

If we start from (12.1) in which $\hat{\mathcal{H}}_\gamma$ is upper Hessenberg we can see that what it implies is that the nth element on the right, i.e., $\varpi_n^*(z)\phi_n(z)$ is written as a linear combination of the elements

$$[\phi_0,\varpi_1(z)\phi_1(z),\ldots,\varpi_{n+1}(z)\phi_{n+1}(z)].$$

These elements generate a space with elements of the form $p_{n+1}(z)/\pi_{n+1}(z)$ with p_{n+1} a polynomial of degree at most $n+1$. It is clear that $\phi_n(z)=\varpi_{n+1}(z)p_n(z)/\pi_{n+1}(z)$ belongs to that space. Hence there must exist an upper Hessenberg matrix $\mathcal{H}$ such that

$$\Phi(\mathcal{I}-\mathit{\Gamma}^*z)\mathcal{H}=\Phi.$$

If the poles are all finite, we could as well use a more explicit π_n^* instead of π_n as denominator of ϕ_n. This means that we denote the poles as γ_j instead of $\overline{\gamma}_j$. Then the previous relation becomes

$$\Phi(z\mathcal{I}-\mathit{\Gamma})\mathcal{H}=\Phi$$

which can be rewritten as

$$z\Phi = \Phi(\mathcal{H}^{-1} + \mathit{\Gamma}).$$

The matrix $\mathcal{H}^{-1}$ is a semiseparable matrix, which is as structured as the Hessenberg matrix. If $\mathcal{H}$ can be written as a product of G_k factors, then the semiseparable $\mathcal{H}^{-1}$ is the product of G_k^{-1} in reverse order.

In the special case of only finite poles and when truncating the problem to a finite dimensional space $\mathcal{L}_n$ in which an inner product with respect to a discrete measure is defined, the previous relation can be truncated to

$$z\Phi_n = \Phi_n(\mathcal{H}_n^{-1} + \mathit{\Gamma}_n) + \phi_{n+1}\mathcal{H}_n^{-1}e_n^T$$

where $e_{n+1} = [0, 0, \ldots, 0, 1]^T$ and $\Phi_n = [\phi_0, \phi_1, \ldots, \phi_n]$. What is implicitly constructed here is the $n+1$-point Szegő quadrature formula since the eigenvalues of $\mathcal{H}_n$ are the zeros of ϕ_{n+1} and the weights are the prescribed weights in the discrete measure. This is the relation given in Theorem 2.7 of [27] where it was obtained in the context of an inverse eigenvalue problem: given the weights and the nodes of the quadrature formula, find the upper Hessenberg matrix (or its inverse) that gives the recurrence for the ORF (plus some extra condition on a unitary matrix formed by the weighted eigenvectors). Obviously the quadrature formula will only be found when one starts from eigenvalues lying on $\mathbb{T}$, in which case the matrix $\mathcal{H}_n$ will be unitary, and it can be written as the product of unitary G_k factors. The idea is however more generally applicable and works not only for ORF on the unit circle but also for ORF on the real line. It is also related to the construction of an orthogonal basis for (rational) Krylov subspaces. For the details we refer to [27] and the papers on which it is based: [12,26] and references therein.

Later, the approach was generalized and both finite and infinite poles were allowed in for example [22]. It combines the previous approach with the one explained in the previous section, to deal with poles at infinity. The general form of the Hessenberg matrix is called an extended Hessenberg (it has blocks that bulge below the subdiagonal) and the G_k factors are called core transformations. See also [17,21,22,1]. For rational Gauss quadrature see [13]. In the context of Krylov subspaces see [24,23].

13 Conclusion

We have systematically developed the basics about orthogonal rational functions on the unit circle whose poles can be freely chosen anywhere in the complex plane as long as they are not on the unit circle $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$. The traditional cases of all poles outside the closed disk and the so-called balanced situations where the poles alternate between inside and outside the disk are included as special instances. The case where all the poles are inside the disk is somewhat less manageable because it is not so easy to deal with the poles at infinity in this formalism, but it has been included anyway. The link between the ORF with all poles outside, all poles inside, and the general case with arbitrary location of the poles is clearly outlined. It is important that previous attempts to consider the general situation were assuming that once some pole $1/\overline{\gamma}$ is chosen, then γ may not appear anywhere in the sequence of remaining poles, which would for example exclude the balanced situation. We have shown how the classical Christoffel-Darboux formula, the recurrence relation, and the relation to rational Szegő formulas can be generalized. Finally we analyzed the matrix representation of the shift operator with respect to the general basis as a matrix Möbius transform of a generalization of a snake-shape matrix, and how the latter can be factorized as a product of elementary unitary 2×2 blocks. It is shown that the most efficient application to compute rational Szegő quadrature formulas is to consider the balanced situation, i.e., the generalization of the CMV representation. In the last section we have made a link with the linear algebra literature where similar, but different algorithms are used in Krylov methods and in inverse eigenvalue problems. As explained in that context, a very similar approach can be taken for ORF on the real line in which case the unitary Hessenberg matrices will be replaced by tridiagonal Jacobi matrices and their generalizations in case poles are introduces in arbitrary order taken from anywhere in the complex plane excluding the real line.

References

[1] J.L. Aurentz, T. Mach, R. Vandebril, and D.S. Watkins. Fast and stable unitary QR algorithm. *Electron. Trans. Numer. Anal.*, 44:327–341, June 2015.

[2] A. Bultheel and M.J. Cantero. A matricial computation of rational quadrature formulas on the unit circle. *Numer. Algorithms*, 52(1):47–68, 2009.

[3] A. Bultheel, M.J. Cantero, and R. Cruz-Barroso. Matrix methods for quadrature on the unit circle. A survey. *J. Comput. Appl. Math.*, 284:78–100, 2015.

[4] A. Bultheel, C. Díaz Mendoza, P. González-Vera, and R. Orive. On the convergence of certain Gauss-type quadrature formulas for unbounded intervals. *Math. Comp.*, 69:721–747, 2000.

[5] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad. *Orthogonal rational functions*, volume 5 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, 1999.

[6] M.J. Cantero, L. Moral, and L. Velázquez. Minimal representations of unitary operators and orthogonal polynomials on the unit circle. *Linear Algebra Appl.*, 408:40–65, 2005.

[7] R. Cruz-Barroso, L. Daruis, P. González-Vera, and O. Njåstad. Sequences of orthogonal Laurent polynomials, bi-orthogonality and quadrature formulas on the unit circle. *J. Comput. Appl. Math.*, 200:424–440, 2007.

[8] R. Cruz-Barroso and S. Delvaux. Orthogonal Laurent polynomials on the unit circle and snake-shaped matrix factorizations. *J. Approx. Theory*, 161(1):65–87, 2009.

[9] R. Cruz-Barroso and P. González-Vera. A Christoffel-Darboux formula and a Favard's theorem for orthogonal Laurent polynomials on the unit circle. *J. Comput. Appl. Math.*, 179(1-2):157–173, 2005.

[10] B. de la Calle Ysern and P. González-Vera. Rational quadrature formulas on the unit circle with arbitrary poles. *Numer. Math.*, 107(4):559–587, 2007.

[11] C. Díaz Mendoza, P. González-Vera, and M. Jiménez Páiz. Strong Stieltjes distributions and orthogonal Laurent polynomials with applications to quadratures and Padé approximation. *Math. Comp.*, 74(252):1843–1870, 2005.

[12] D. Fasino and L. Gemignani. Direct and inverse eigenvalue problems for diagonal-plus-semiseparable matrices. *Numer. Algorithms*, 34(2–4):313–324, 2003.

[13] D. Fasino and L. Gemignani. Structured eigenvalue problems for rational Gauss quadrature. *Numer. Algorithms*, 45(1–4):195–204, 2007.

[14] B. Fritzsche, B. Kirstein, and A. Lasarow. Orthogonal rational matrix-valued functions on the unit circle. *Math. Nachr.*, 278(5):525–553, 2005.

[15] B. Fritzsche, B. Kirstein, and A. Lasarow. Orthogonal rational matrix-valued functions on the unit circle: Recurrence relations and a Favard-type theorem. *Math. Nachr.*, 279:513–542, 2006.

[16] B. Fritzsche, B. Kirstein, and A. Lasarow. Para-orthogonal rational matrix-valued functions on the unit circle. *Oper. Matrices*, 6(4):631–680, 2012.

[17] L. Gemignani. A unitary Hessenberg QR-based algorithm via semiseparable matrices. *J. Comput. Appl. Math.*, 184(2):505–517, 2005.

[18] W.B. Jones, O. Njåstad, and W.J. Thron. Two-point Padé expansions for a family of analytic functions. *J. Comput. Appl. Math.*, 9:105–124, 1983.

[19] W.B. Jones, O. Njåstad, and W.J. Thron. Orthogonal Laurent polynomials and the strong Hamburger moment problem. *J. Math. Anal. Appl.*, 98:528–554, 1984.

[20] A. Lasarow. *Aufbau einer Szegő-Theorie für rationale Matrixfunktionen.* PhD thesis, Universität Leipzig, Fak. Mathematik Informatik, 2000.

[21] T. Mach, M. Van Barel, and R. vANdebril. Inverse eigenvalue problems linked to rational Arnoldi, and rational, (non)symmetric Lanczos. Technical Report TW629, Dept. Computer Science, KU Leuven, June 2013.

[22] T. Mach, M. Van Barel, and R. Vandebril. Inverse eigenvalue problems for extended Hessenberg and extended tridiagonal matrices. *J. Comput. Appl. Math.*, 272:377–398, 2014.

[23] C. Mertens. *Short Recurrence Relations for (Extended) Krylov Subspaces*. PhD thesis, Numerical Analysis and Applied Mathematics Section, Department of Computer Science, Faculty of Engineering Science, April 2016. Vandebril, Raf and Van Barel, Marc (supervisors).

[24] C. Mertens and R. Vandebril. Short recurrences for computing extended Krylov bases for Hermitian and unitary matrices. *Numer. Math.*, 131(2):303–328, 2015.

[25] W.J. Thron. L-polynomials orthogonal on the unit circle. In A.M. Cuyt, editor, *Nonlinear Numerical Methods and Rational Approximation*, pages 271–278. Kluwer, 1988.

[26] M. Van Barel, D. Fasino, L. Gemignani, and N. Mastronardi. Orthogonal rational functions and diagonal-plus-semiseparable matrices. In F.T. Luk, editor, *Advanced Signal Processing Algorithms, Architectures, and Implementations XII*, volume 4791 of *Proceedings of SPIE*, pages 162–170. SPIE, 2002.

[27] M. Van Barel, D. Fasino, L. Gemignani, and N. Mastronardi. Orthogonal rational functions and structured matrices. *SIAM J. Matrix Anal. Appl.*, 26(3):810–829, 2005.

[28] L. Velázquez. Spectral methods for orthogonal rational functions. *J. Funct. Anal.*, 254(4):954–986, 2008.